

Bioinformatische Untersuchungen zur Evolution der Prokaryoten und Plastiden

Von der Gemeinsamen Naturwissenschaftlichen Fakultät
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig

zur Erlangung des Grades einer
Doktorin der Naturwissenschaften
(Dr.rer.nat.)

genehmigte
Dissertation

von Sabine Cornelsen
aus Braunschweig

1.Referent: Prof. Dr. W. F. Martin

2.Referent: Prof. Dr. R. Cerff

eingereicht am: 4.2.2002

mündliche Prüfung (Disputation) am: 28.3.2002

2002

(Druckjahr)

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Naturwissenschaftlichen Fakultät, vertreten durch Prof. Dr. W. F. Martin, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Martin W., Hansmann S. Phylogeny of 33 ribosomal and six other proteins encoded in an ancient, conserved gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis. *International Journal of Systematic and Evolutionary Microbiology*, 50:1655-1663, 2000.

Martin W., Stoebe B., Goremykin V., Hansmann S., Hasegawa M., Kowallik V. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*, 393:162-165, 1998.

Danksagungen

Im Laufe meiner Promotion habe ich von vielen Seiten Unterstützung erfahren, ohne die die Fertigstellung dieser Arbeit nicht möglich gewesen wäre. Ich möchte nun an dieser Stelle denen danken, die mir in den letzten Jahren bei dieser Arbeit geholfen haben.

bei **Professor Bill Martin**, der mich schon während meines Studiums für die Problematik phylogenetischer Algorithmen begeistern konnte und mich seit dieser Zeit unterstützt und gefördert hat. So ermöglichte er mir Mitarbeit an einigen wissenschaftlichen Arbeiten sowie den Forschungsaufenthalt in Neuseeland. Die Zusammenarbeit, in der er fortwährendes Interesse an meinen Ergebnissen zeigte, stellten die Grundlage zur Fertigstellung dieser Dissertation dar.

bei **Professor David Penny** und **Dr Peter Lockhart** (Massey University, Palmerston North, New Zealand), die mir die Grundideen der phylogenetischen Spektralanalyse näher gebracht und erklärt haben

bei **Professor Cerff**, der mir eine zeitweilige Beschäftigung und den Zugang zu einem Arbeitsplatz an seinen Institut ermöglichte und sich sofort bereit erklärte, diese Arbeit als Koreferent zu begutachten.

bei den Mitarbeitern am **Institut für Genetik** (TU Braunschweig), am **Institut für Botanik III** (Universität Düsseldorf) und am **Institute for Molecular BioScience** (Massey University, Palmerston North, New Zealand), für die jeweils freundliche Aufnahme und manches nettes und hilfreiches Gespräch.

bei **Dr Ulrich Nowitzki**, der mir mit der Durchsicht und seinen kritischen Anmerkungen bei der Überarbeitung dieser Arbeit sehr geholfen hat.

bei der **Heinrich Böll Stiftung**, die mich seit dem 1.4.1999 durch ein Promotionsstipendium finanziell gefördert hat.

bei **Tim Wittrock**, der mir häufig weiterhalf, wenn mir der richtige UNIX Befehl nicht einfiel oder ich Probleme beim Programmieren hatte.

bei meinen **Freunden** und meiner **Familie** für Ihr Interesse am Fortgang meiner Arbeit und Ihre vielen hilfreichen Unterstützungen.

bei meinen **Eltern**, die mir während meines Studiums und meiner Promotion durch Ihre liebe hilfreiche Unterstützung stets zur Seite gestanden haben.

bei meinem Mann **Karsten**, der mich mit viel Liebe in den Höhen und Tiefen während dieser Arbeit begleitet hat. Er war bei Problemstellungen stets ein hilfreicher Ansprechpartner und konstruktiver Kritiker.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Molekulare Marker und phylogenetische Algorithmen	3
1.2	Evolution der Bakterien	4
1.3	Evolution der Plastiden	6
1.4	Zielsetzung	8
2	Material und Methoden	9
2.1	Algorithmen und Methoden der Phylogenetik	9
2.1.1	Die Neighbor-Joining Methode	10
2.1.2	Maximum Likelihood Analyse mit dem Programm PROTML	11
2.1.3	PROTDET - Vom Spektrum aller Teilbäume zum Ab- stammungsbaum	14
2.1.4	RASA - Relative Apparent Synapmophy Analysis	17
2.1.5	Die Variabilität von Sequenzpositionen als Maß für phy- logenetische Information	19
2.2	Die untersuchten Datensätze	23
2.2.1	Der Bakteriendatensatz	23
2.2.2	Der Plastidendatensatz	28
3	Ergebnisse	35
3.1	Untersuchungsschritte	35
3.2	Auswertung der Bakteriendatensätze	36
3.2.1	Der Archae- und Eubakteriendatensatz	36
3.2.2	Der Eubakteriendatensatz	39
3.3	Auswertung des Plastidendatensatzes	49
3.3.1	Qualitativer Vergleich der Topologien	52
4	Diskussion	59
4.1	Die angewandten phylogenetischen Methoden im Vergleich	60
4.1.1	Die RASA Untersuchungen	60
4.1.2	PROTML contra PROTDET	60
4.2	Untersuchung der Archae- und Eubakterien	61
4.3	Die Untersuchung des Plastidendatensatzes	62
4.4	Bewertung	64
5	Zusammenfassung und Ausblick	65

A Benutzerhandbuch	67
A.1 Eingabeformat des Alignments	67
A.2 Das Programm SORTAL	68
A.3 Das Programm PBL	68
A.4 Das Programm RASA	74
B Implementierung der Programme	76
C Der Eubakteriendatensatz	78
D Der Plastidendatensatz	112

Abbildungsverzeichnis

1.1	Phylogenetischer Abstammungsbaum nach Haeckel,1866	2
1.2	16s rRNA Baum der Eubakterien	6
1.3	16s rRNA Baum der Purpurbakterien	6
1.4	Stammbaum der Algenarten und Pflanzen	7
2.1	Die Neighbor Joining Methode	11
2.2	ProtML: Beispielbaum mit 5 Blättern	12
2.3	Das Prinzip von <i>Local Rearrangement</i>	13
2.4	PROTDET: Beispielbaum mit 5 Blättern	14
2.5	Beispielalignment vor dem Sortieren	21
2.6	Beispielalignment nach dem Sortieren	22
2.7	Die konserviertesten Abschnitte des Wächtershäuser Clusters . .	27
3.1	RASA-Analyse Archae- und Eubakterien	37
3.2	RASA-Analyse der Variabilitätsgruppen 2 bis 11 des Archae- und Eubakterien Alignments	38
3.3	RASA-Analyse Eubakterien	40
3.4	RASA-Analyse der Variabilitätsgruppen 2 bis 11 des Eubakterien Alignments	41
3.5	RASA-Analyse der Variabilitätsgruppen 12 bis 20 des Eubakte- rien Alignments	42
3.6	Phylogenetische Bäume des Eubakteriendatensatzes	43
3.7	Die PROTDET Frequenzen der Splits bezüglich des gesamten und des RASA Eubakterien Alignmnents	48
3.8	RASA-Analyse Plastidendatensatz	50
3.9	RASA-Analyse der Variabilitätsgruppen 2 bis 11 des Plastiden Alignments	55
3.10	RASA-Analyse der Variabilitätsgruppen 12 bis 14 des Plastiden Alignments	56
3.11	Phylogenetische Bäume des Plastidendatensatz	57
3.12	Die PROTDET Frequenzen der Splits bezüglich des gesamten und des RASA Plastiden Alignmnents	58
4.1	Zuordnung der Eubakterien zum 16s rRNA Baum	62
4.2	Zuordnung der untersuchten Organismen zum Stammbaum der Algen und Landpflanzen	63
B.1	Datenstruktur zum Verwalten der Alignmentsinformationen . . .	77

Tabellenverzeichnis

2.1	Die Splits, Bipartitionen und Teilmengen eines 5 blättrigen Baumes	15
2.2	Bakteriendatensatz	25
2.3	Übersicht der einzelnen Bakterienalignments	26
2.4	Plastidendatensatz	28
2.5	Kodierung der Chloroplastenproteine	33
2.6	Plastidenalignments	34
3.1	Übersicht der Variabilitätsgruppen des Archae-/ Eubakterienalignments	36
3.2	Übersicht der Variabilitätsgruppen des Eubakterienalignments . .	39
3.3	Die Local Bootstrapwerte der ML Topologie des EubakterienAlignments	46
3.4	Die Splits der Eubakterienstammbäume	47
3.5	Übersicht der Variabilitätsgruppen des Plastidenalignments . . .	50
3.6	Die Local Bootstrapwerte der ML Topologie der PlastidenAlignments	52
3.7	Die Splits der Plastidenstammbäume	53

Kapitel 1

Einleitung

Es ist ein langgehegter Wunsch der Biologen alle Lebewesen zu klassifizieren und sie dann in hierarchische Strukturen einzuordnen. Aber durch Darwin (Darwin (1859)) setzte sich die Abstammungslehre und die natürliche Klassifikation durch, wonach alle Lebewesen miteinander verwandt sind und auf einen gemeinsamen Vorfahren zurückgeführt werden können.

Bis Mitte des 20. Jahrhunderts beruhten biologische Klassifizierungssysteme auf sichtbare, morphologische Merkmale der Form von Zellen, Organismen und Fossilien. So teilte Ernst Haeckel 1866, ausgehend von Darwins Theorie, die Lebewesen in die drei Reiche der Tiere, der Pflanzen und der Protisten auf. Diese faßte er durch die damals bekannten Gattungen und Arten von Pflanzen und Tieren unter dem Gesichtspunkt ihrer möglichen Entwicklung in einem phylogenetischen Baum zusammen (vgl. Haeckel (1866)). Besonders problematisch war dabei die natürliche Klassifizierung der Einzeller, weil sie besonders arm an auswertbaren morphologischen Merkmalen waren.

So umfaßte Haeckels Gruppe „Protisten“ neben den Organismen, die heute als einzellige Eukaryoten (Zellen mit echtem Kern) bezeichnet werden, auch die von ihm als „Moneres“ gekennzeichneten Organismen, welche heute als Prokaryoten bezeichnet werden (vgl. Abbildung 1.1). Die Einteilung der Einzeller in Prokaryoten und Eukaryoten erfolgte erst ca. 80 Jahre nach Haeckels Veröffentlichung durch van Niel (van Niel (1946)).

Jedoch blieben alle Versuche, eine natürliche d.h. in der Abstammung begründete Klassifikation sowohl der Prokaryoten als auch der einzelligen Eukaryoten (heute als Protisten im engeren Sinne bezeichnet) zu erstellen, bis in die Gegenwart problematisch. Der Grund hierfür war der gleiche mit dem sich auch Haeckel konfrontiert sah, nämlich das Fehlen von abgestuften Merkmalsreihen, welche auf eine eindeutige Verwandtschaft schließen lassen könnten.

Die Probleme in der Erstellung einer natürlichen Systematik der Einzeller werden in Corliss (1994) in sehr übersichtlicher Form dargestellt. Dieser Artikel beinhaltet auch eine Übersicht mehrerer Systeme des vergangenen

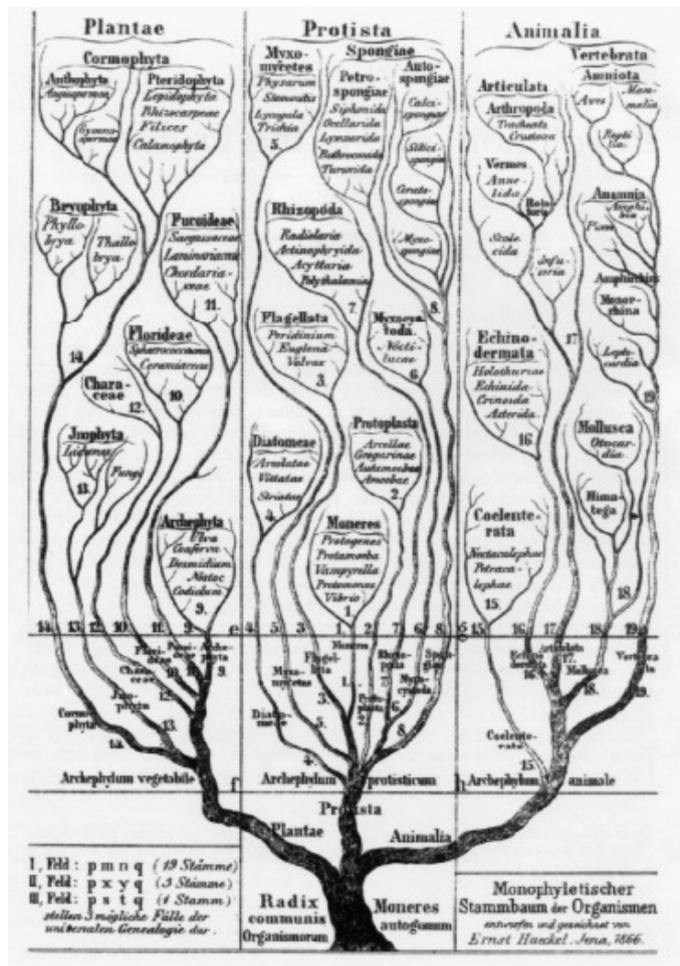


Abbildung 1.1: Phylogenetischer Abstammungsbaum nach Haeckel, 1866 (Haeckel (1866))

Jahrhunderts. Corliss Klassifizierung unterscheidet sich aber auch grundlegend von der in Cavalier-Smith (Cavalier-Smith (1998)) vorgeschlagenen hinsichtlich der Gruppeneinteilung und insbesondere auch in der Namensgebung.

Einen möglichen Ausweg aus der Problematik der Klassifizierung von Einzellern zeigten Zuckerkandl und Pauling (Zuckerkandl and Pauling (1965)) auf. Sie schlugen vor, daß die Sequenzen von Proteinen unter Umständen dazu genutzt werden könnten, vergleichende Studien zur Evolution durchzuführen. So wurden in der Zeit zwischen 1970 und 1980 zum Zweck der Aufklärung der Stammesgeschichte der Prokaryoten sehr viele Sequenzen für das Protein Cytochrome c erstellt und verglichen. Dabei wurde jedoch das Problem ersichtlich, daß Gene für Cytochrome c zwischen unterschiedlichen Bakteriengruppen lateral tranferiert werden konnten. (Dickerson (1980)).

Als Standardmolekül für die Untersuchung phylogenetischer Beziehungen

innerhalb der Gruppe der Einzeller setzte sich nicht das Cytochrome c sondern die kleine Untereinheit der ribosonalen RNA, die 16s rRNA durch. Durch die Studie dieses Moleküls konnten zum Beispiel die Archaeobakterien entdeckt werden (Woese and Fox (1977)). Diese unterschieden sich in vielen biochemischen Merkmalen grundlegend von den übrigen Prokaryoten, den Eubakterien (Woese (1987)). Die Zweiteilung der Prokaryoten wird von vielen Autoren heute als die tiefste, d.h. die evolutionär älteste, Einteilung in der natürlichen Klassifizierung der Organismen angesehen (Doolittle (1999)).

Obwohl die ribosonale RNA die eindeutige Zweiteilung der Prokaryoten in Archaeobakterien und Eubakterien widerspiegelt und diverse Gruppierungen innerhalb dieser Gruppe erkennen läßt, bleiben mit dem Maß der 16s rRNA Sequenzähnlichkeit die Beziehungen der Prokaryotengruppe untereinander weitestgehend unaufgelöst (Pace (1997)). Aber trotz der heutigen verfügbaren Daten, sogar aus ganzen Genomen, hat sich kein Molekül, Gen oder Gruppe von Genen als akzeptierte Alternative zu rRNA präsentieren können.

1.1 Molekulare Marker und phylogenetische Algorithmen

Die Fragen nach geeigneten Markern und Methoden, die aus dem Sequenzvergleich Rückschlüsse auf evolutionäre Prozesse ziehen können, beschäftigen seit Zuckerkandl und Paulings Artikel bis heute die Systematiker.

Die Güte der evolutionären Informationen, die sich aus einem Sequenzvergleich ergeben, hängen maßgeblich von der Wahl des molekularen Markers ab. Ein für die Evolutionsforschung universell anwendbares Gen sollte dabei im Prinzip die folgenden Anforderungen erfüllen (Woese (1987)).

- Vorhandensein in allen Organismen
- Änderungen der Sequenzpositionen sollten zufällig verteilt erscheinen
- Sequenzgröße sollte ausreichend erscheinen, um hinreichend Informationen zu beinhalten

Bei den ersten phylogenetischen Sequenzvergleichen war es zudem auch wichtig, daß sich die Gene schnell sequenzieren ließen. Ein weiteres zu berücksichtigendes Problem stellt der laterale bzw. horizontale Gentransfer dar (Figge and Cerff (2001)). Dieser bezeichnet die Bewegung von Genmaterial zwischen entfernt verwandten Organismen. Ist ein Gen davon betroffen, läßt sich aus dem daraus erstellten phylogenetischen Baum zwar eine Genphylogenie aber nur beschränkt eine Organismenphylogenie ableiten.

In dem Zeitraum von 1965 bis 1977 beruhten die molekularen Untersuchungen vorwiegend auf Proteinen, speziell Ferredoxin und Cytochrome (Doolittle (1999)). Mitte der 70-iger Jahre begannen Woese und Mitarbeiter mit Untersuchungen von Prokaryoten anhand von 16s rRNA (Woese and Fox (1977)), Woese

(1987)). Diese erwies gegenüber den bisherigen Markern einige Vorteile. Neben der leichten Sequenzierbarkeit und der Verbreitung in prokaryotischen und eukaryotischen Zellen, weisen 16s rRNAs schnell und langsam evolvierende Abschnitte sowie konservierte Strukturen auf. Zudem verhalten die „alte“ und essentielle Funktion innerhalb des Zellhaushaltes sowie die Interaktion mit zahlreichen RNAs und Proteinen den sRNA Sequenzen zu dem Status der am häufigsten verwendeten molekularen Marker.

Zur Zeit steigt die Anzahl sequenzierter Proteine bzw. ganzer Genome durch neue technische Verfahren rasant an. Die Fülle an Daten, die heute in Datenbanken weltweit zur Verfügung stehen, ermöglichen es nicht nur einzelne Proteine oder rRNA Sequenzen sondern ganze Genome zu vergleichen. Hierbei werden solche Proteine berücksichtigt, die in allen zu untersuchenden Genomen vorhanden sind. Dies hat den Vorteil nicht nur den evolutionären Einfluß auf ein Protein oder eine rRNA Sequenz zu beobachten, sondern durch Berücksichtigung aller Homologen eine Aussage über die evolutionären Vorgänge ganzer Genome zu gewinnen.

Neben der Wahl eines geeigneten molekularen Markers stellt auch die Wahl einer geeigneten Methode zum Sequenzvergleich ebenfalls einen entscheidenden Faktor bei phylogenetischen Untersuchungen dar. Die zahlreichen entwickelten Methoden versuchen anhand unterschiedlicher Kriterien aus dem Sequenzvergleich Rückschlüsse über die Veränderungen der Sequenzen im Laufe der Evolution zu ziehen. Aber mit robusten Methoden konnten grundlegende evolutionäre Erkenntnisse gewonnen werden. Mit robusten Methoden bezeichnet man u.a. solche Methoden, die in Simulationstests gute Ergebnisse ergeben haben (vgl. Nei and Saitu (1987), Nei (1996)).

Für diese Tests werden aus jeweils einer Sequenz anhand eines Modellabstammungsbaumes und statistischer Verteilungsfunktionen zum Verändern einzelner Sequenzpositionen Datensätze erstellt, die aus zu vergleichenden Sequenzen bestehen. Die Robustheit einer Methode mißt sich bei solchen Tests daran, inwieweit die richtige Topologie des Modellabstammungsbaumes mit der gegebenen Methode rekonstruiert werden konnte.

Die im folgenden beschriebenen evolutionären Beziehungen der Bakterien und der Plastiden beruhen auf phylogenetischen Untersuchungen molekularer Marker wie 16s rRNA.

1.2 Evolution der Bakterien

Die Prokaryoten umfassen alle Bakterienarten. Ihre Zellen besitzen im Gegensatz zu Eukaryoten keinen von einer Membran umgebenden Kern und keine Organellen. Das Erbmateriale, die DNA, liegt in der Regel als ringförmig geschlossenes Chromosom im Zellplasma vor. Bei Eukaryoten ist die DNA im Zellkern in ein oder mehreren linearen Chromosomen zu finden.

Eine Klassifizierung der Bakterien stellte sich zunächst auf Grund der geringen morphologischen Unterschiede als nahezu unmöglich dar. Die Einteilung, ob prokaryotische Zellen nach der Methode von Christian Gram anfärbbar sind, führte im Jahre 1884 zu einer ersten sinnvollen phylogenetischen Aufteilung in Gram-positive und Gram-negative Bakterien (Gram (1884)). Bis in die 70-iger Jahre des 20. Jahrhunderts sollte es auch keine weiteren elementaren Klassifizierungsansätze geben. Versuche die evolutionäre Beziehungen der Prokaryoten zu erfassen, ergaben wie zum Beispiel von C.B. van Niel vielversprechende Ansätze bezüglich der Klassifikation, führten aber in den folgenden Jahren zu keinen nachhaltigen Resultaten (Stainer and van Niel (1941), van Niel (1946)).

Erst Woese und Mitarbeitern gelang durch Sequenzierung und Vergleich von 16s rRNA eine Klassifizierung, die das Erstellen erster phylogenetischer Stammbäume ermöglichte (Woese (1987), Gupta (1998)). Woeses Untersuchungen führten zu einer grundlegenden Entdeckung. Anhand der Sequenzvergleiche ließen sich zwei Arten von Bakterien erkennen, die sich in einem solchen Maß voneinander unterschieden, daß die Anschauung von der Zweiteilung aller lebenden Organismen in Prokaryoten und Eukaryoten von der Einteilung in die drei Gruppen Eubakterien, Archaeobakterien und Eukaryoten abgelöst wurde (Woese and Fox (1977), Gupta (1998)). Die Klassifizierung anhand von 16s rRNA, wie sie von Woese begonnen wurde, und den daraus resultierenden Stammbäume, stellen bis heute die Grundlage der bakteriellen Evolution dar, die im folgenden kurz erläutert werden soll.

Archaeobakterien kommen in sehr extremen Lebensräumen wie aktive Vulkangebiete oder sehr salzhaltigen Umgebungen vor. Zum gegenwärtigen Zeitpunkt teilen sich die Archaeobakterien in die drei Gruppen methanogene (methanbildende) Bakterien, halophile (salzliebende) Bakterien und thermo-acidophile (wärme-/säureliebende) Bakterien auf.

Bei den Eubakterien ist Klassifizierung weitaus komplexer. Der Stammbaum in Abbildung 1.2 zeigt einige Eubakteriengruppen (vgl. auch Schlegel (1992)).

Es ist gut zu erkennen, daß das Unterscheidungsmerkmal der Anfärbbarkeit nach Gram im Falle der Gram-positiven Bakterien ein Kriterium einer eigenen Klasse darstellt. Sie sind in diesem Baum gemeinsam mit den Cyanobakterien und den Purpur-/ bzw. Proteobakterien an einem Ast zu finden. Wie in Abbildung 1.3 zu sehen ist, lassen sich die Purpurbakterien wiederum in die vier Gruppen α , β , γ und δ/ϵ unterteilen.

Eine erste Untersuchung ganzer prokaryotischer Genome behandelte Wächtershäuser in Wächtershäuser (1998). Er untersuchte die Anordnung der Gene in den Genomen von 12 Eubakterien und 7 Archaeobakterien und entdeckte ein konserviertes Muster bzw. Cluster. In diesem Cluster lokalisierte er homologe Gene für die Archaeobakterien, die Eubakterien sowie für alle Prokaryoten. Anhand der Homologen und deren Anordnung im Cluster zog er

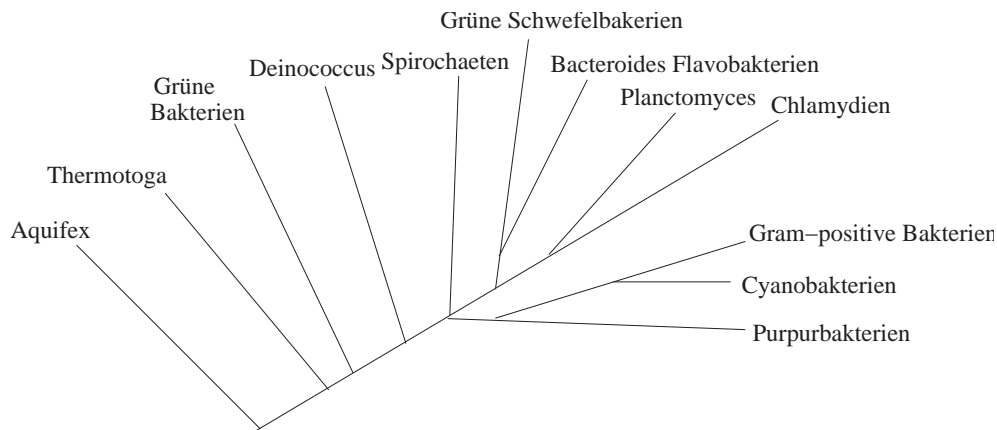


Abbildung 1.2: Dieser Abstammungsbaum der Eubakterien beruht auf Untersuchungen von 16s rRNA angelehnt an das Lehrbuch von Schlegel (Schlegel (1992)).

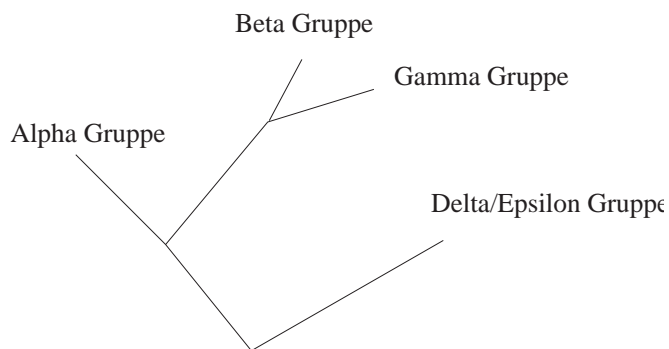


Abbildung 1.3: Der Einteilung der Purpurbakterien in die α , β , γ und δ/ϵ Gruppen liegt wiederum der Vergleich von 16s rRNA zugrunde.

Rückschlüsse über das mögliche Aussehen des Genoms einer hypothetischen Urzelle.

Der Artikel bot den Ausgangspunkt zur Untersuchung der bakteriellen Evolution in dieser Arbeit.

1.3 Evolution der Plastiden

Eukaryotische Zellen weisen im Gegensatz zu Prokaryoten neben dem Zellkern Organellen wie Mitochondrien und im Falle von Algen/Pflanzen auch Chloroplasten auf. Neben dem Kern enthalten auch die Organellen DNA, die wie bei Bakterien in einem ringförmig geschlossenen Chromosom vorliegt. Organellen, die von einer Doppelmembran umgeben sind, d.h. Chloroplasten und Mitochondrien, besitzen so auch meistens eine eingeschränkte Möglichkeit zur Proteinsynthese. Außerdem gibt es Übereinstimmungen zwischen der Organell-

und der Prokaryonten Membran. Die Ähnlichkeiten zwischen Organellen und Prokaryonten unterlegen die sogenannte Endosymbiontenhypothese (Goksøyr (1967), Sagan (1967)).

Diese besagt, daß sich Mitochondrien von aeroben Bakterien, wahrscheinlich Proteobakterien, und die Chloroplasten von Cyanobakterien herleiten lassen. Es wird angenommen, daß die Bakterien nach und nach in die Eucyten, einzellige Vorgänger der Eukaryoten, eingedrungen und eine Endosymbiose, eine Symbiose innerhalb der Eucyte, eingegangen sind (für eine Übersicht siehe Martin *et al.* (2001)).

Phylogenetische Untersuchungen durch den Sequenzvergleich einzelner Proteine oder ganzer Genome von Prokaryoten gegenüber einzelnen Genen bzw. den Genomen von Mitochondrien oder Chloroplasten führen zu aussagekräftigeren Aussagen als bei der Untersuchungen von Prokaryoten. Der Grund dafür ist, daß die Gene in den Mitochondrien- und Chloroplastengenomen in einer sehr konservierten Form vorliegen. Auch wenn im evolutionären Prozeß einige Gene in den Genom verlorengegangen sind, waren diese nicht vom lateralen Gentransfer betroffen.

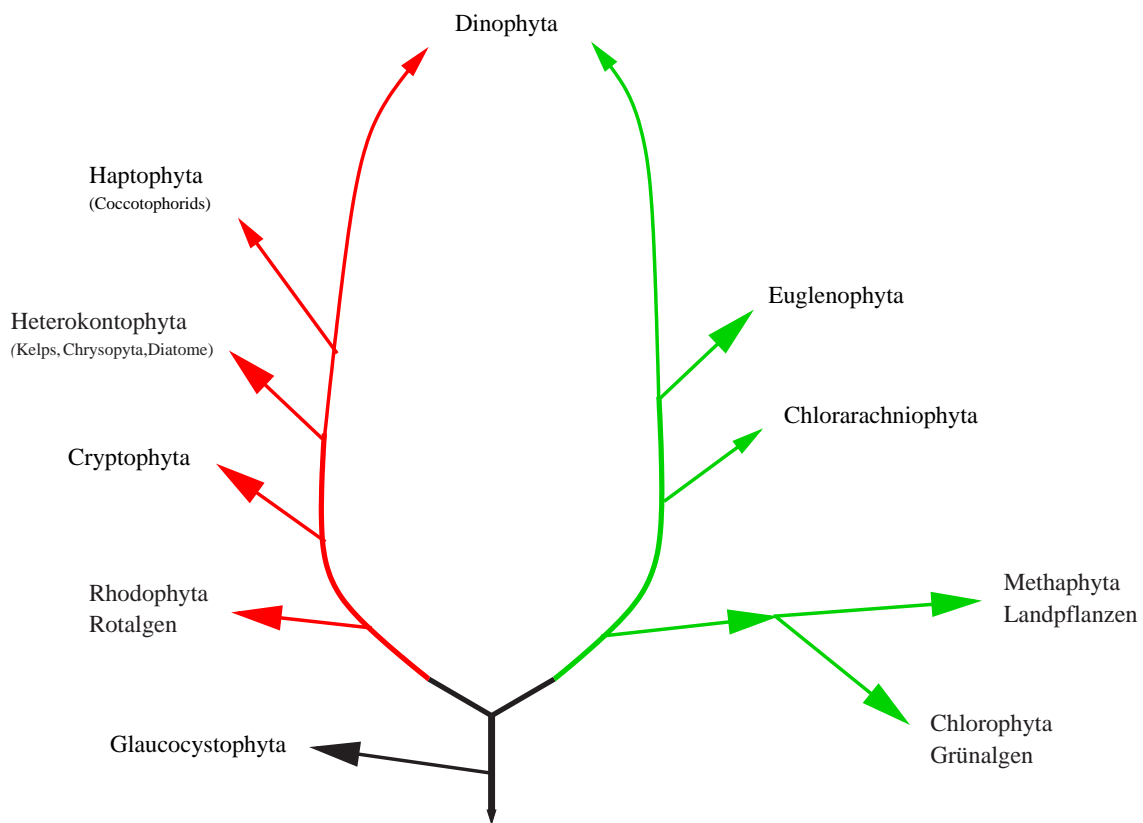


Abbildung 1.4: Dieser Abstammungsbaum der Algen und Pflanzen lehnt sich die Darstellung in Delwiche (1999) an

In Abbildung 1.4 zeigt die aus bisherigen Untersuchungen und Erkenntnissen resultierende Phylogenie der Algen und Landpflanzen angelehnt an den Artikel von Delwiche (Delwiche (1999)). Mit Ausnahme der Glaucocystophyta lassen sich die Algen/Pflanzen in zwei große Gruppe aufteilen. Der Ast der Rot-/Braunalgen sowie Verwandter ist rot, der Ast der Landpflanzen, Grünalgen und Verwandter ist grün hervorgehoben.

Die Glaucocystophyta können als ursprünglichste Eukaryotengruppe angesehen werden, die mit Cyanobakterien eine Symbiose eingegangen sind. Die Untersuchungen bestätigen, daß lediglich die Rotalgen (Rhodophyta) und Grünalgen (Chlorophyta) durch Umwandlung von Algenarten der Glaucocystophyta entstanden sind (primäre Endosymbiose). Die Gruppe der höheren Pflanzen stellt eine hoch spezialisierten Ast der Chlorophyta dar.

Weiterhin werden die Vermutungen bekräftigt, daß die Chloroplasten vieler Algenarten aus der Aufnahme primitiver Algenarten wie Glaucocystophyta bzw. primitiver Rot-/Grünalgen in Eukaryoten hervorgegangen sind (sekundäre Endosymbiose). Bei der Gruppe der Dinophyta läßt sich sogar eine tertiäre Endosymbiose feststellen. Dieser Vorgang beschreibt die Aufnahme einer durch sekundär Endosymbiose entstandenen Zelle in eine eukaryotische.

1.4 Zielsetzung

Phylogenetische Bäume zur Untersuchung der Evolution der Prokaryoten oder der Plastiden, wie sie in zahlreichen Lehrbüchern und Veröffentlichungen zu finden sind, beruhen meist auf phylogenetischen Untersuchungen von rRNA oder einzelnen Proteinen. Solche phylogenetischen Bäume können allerdings nur die evolutionären Beziehungen bezüglich eines Gens wiedergeben, die nicht unbedingt dem gesamten evolutionären Prozeß entsprechen müssen.

Bei den Untersuchungen dieser Arbeit sollen nicht nur einzelne Proteine als Grundlage dienen, sondern möglichst viele Gene, die in den entsprechenden Genomen vorhanden sind. Die analysierten Daten der Bakterien- bzw. der Plastidengenome setzen sich nur aus solchen Homologen zusammen, deren Sequenzvergleich nicht auf einen lateralen Gentransfer schließen läßt.

Zum einen ist es Ziel dieser Arbeit mit unterschiedlichen Methoden zu überprüfen, ob Daten aus sequenzierten prokaryotischen Genomen und aus sequenzierten Plastidengenomen hinreichend phylogenetische Informationen enthalten und inwieweit die erstellten Bäume Differenzen zu den bisher bekannten Topologien aufweisen. Zum anderen sollen Möglichkeiten analysiert werden, informative Abschnitte in den Daten zu lokalisieren, um diese gezielt als Grundlage zur Rekonstruktion phylogenetischer Bäume heranziehen zu können.

Kapitel 2

Material und Methoden

Die untersuchten Datensätze beinhalten homologe Aminosäuresequenzen von ausgewählten Proteinen. Die 20 verschiedenen Aminosäuren werden durch die Buchstaben $\{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$ repräsentiert. Mit Programmen CLUSTALW Thompson *et al.* (1994) und PILEUP GCG (1994) werden die Sequenzen nach übereinstimmenden Mustern in der Anordnung der Aminosäuren durchsucht. Übereinstimmungen werden einander zugeordnet und die Sequenzen bei fehlenden Mustern durch sogenannte „gaps“ an entsprechenden Stellen ergänzt. Der entstandene Datensatz, das Alignment, enthält somit Sequenzen der gleichen Länge.

Solche Alignments bilden den Ausgangspunkt der phylogenetischen Untersuchungen, wie sie im ersten Abschnitt des Kapitels vorgestellt werden. Der zweite Abschnitt beschreibt den Aufbau der zwei Datensätze, deren Untersuchung Gegenstand der folgenden Kapitel sein werden.

2.1 Algorithmen und Methoden der Phylogenetik

Ein phylogenetischer Baum soll die evolutionären Beziehungen zwischen verschiedenen Organismen darstellen. In solchen Abstammungsbäumen sind diese Organismen durch die äußeren Knoten repräsentiert. Diese werden auch als Blätter oder OTUs (operational taxonomical units) bezeichnet. Die inneren Knoten stehen für bereits ausgestorbene Vorgängerorganismen.

Zur Rekonstruktion phylogenetischer Bäume aus DNA oder Protein Alignments wurden zahlreiche Modelle entwickelt. Die Kriterien lassen sich prinzipiell in unterschiedliche Kategorien einteilen. Die drei verschiedenen phylogenetischen Rekonstruktionsmethoden, die Gegenstand der Analysen dieser Arbeit sind, lassen sich den folgenden drei Kategorien, zuordnen.

- Distanzmethoden

Für jedes Sequenzpaar wird eine Distanz berechnet, die auf dem paarweisen Vergleich der Sequenzen beruht. Die Distanz soll ein Maß für die durchschnittliche Substitutionsrate des Sequenzpaares darstellen. Das

bedeutet, je geringer die Distanz desto geringer ist die Substitutionsrate und umso höher die Ähnlichkeit des Sequenzpaares.

Anhand der Distanzen werden nun Bäume erstellt deren Gesamtdistanzwert möglichst gering sein soll. Mit der Neighbor Joining Methode (vgl. Nei and Saitu (1987)) wird einer der etabliertesten Vertreter dieser Methoden unter 2.1.1 vorgestellt.

- **Maximum Likelihood Methoden**

Die Grundidee der Maximum Likelihood (ML) Methoden ist es, den Baum mit der höchsten Wahrscheinlichkeit, dem höchsten Likelihood zu finden. Die Berechnung des Likelihood eines Baumes (vgl. Abschnitt 2.1.2) ermöglicht es zum Beispiel mehrere Bäume miteinander zu vergleichen und daraus den besten, den mit dem höchsten Likelihood, auszuwählen. Mit den ML Methoden lassen sich aber auch die Bäume schrittweise aufbauen, indem die wahrscheinlichsten Teilbäume bestimmt werden, aus denen sich später der gesamte ML Baum zusammensetzt. Das Vorgehen mit dem ML Programm PROTML Adachi and Hasegawa (1996) wird in Abschnitt 2.1.2 beschrieben.

- **Spektralanalysen**

Bei Spektralanalysen werden die Teilbäume betrachtet, aus denen sich ein Baum zusammensetzt. Der erste Schritt besteht darin, ein Spektrum über alle möglichen Teilbäume bzw. Splits eines Datensatzes zu bestimmen. Dieses soll für jeden Split wiedergeben, wie groß der Anteil der Positionen des Datensatzes ist, die nach einem bestimmten Kriterium einen Split widerspiegeln. Der resultierende Baum besteht aus den besten, sich nicht ausschließenden Splits. Das in Abschnitt 2.1.3 vorgestellte Programm PROTDET führt eine Spektralanalyse eines Datensatzes durch.

In diesem Abschnitt werden nun die für die Untersuchungen relevanten Methoden und die entsprechend verwendeten Programme vorgestellt.

2.1.1 Die Neighbor-Joining Methode

Die Neighbor Joining (NJ) Methode (Nei and Saitu (1987)) bestimmt anhand einer Distanzmatrix einen phylogenetischen Abstammungsbaum. Die Distanzmatrix enthält für jedes Sequenzpaar eine Distanz, die der durchschnittlichen Anzahl von Substitutionen des Sequenzpaares entsprechen soll. Die Distanzen werden anhand des Vergleichs von je zwei Sequenzen bestimmt. Die Berechnungsmethoden zur Ermittlung der Distanz, auf die hier nicht näher eingegangen werden soll, verwenden dabei unterschiedliche statistische Verteilungsmodelle und Konstantenmatrizen, um Rück- und Mehrfachsubstitutionen berücksichtigen zu können (s. Graur (1999)).

Ausgangspunkt des Verfahrens ist eine Sterntopologie, wie sie Abbildung 2.1 a) zu sehen ist. Der erste Schritt besteht darin, einen inneren Knoten für ein

Sequenzpaar so einzufügen, daß die Gesamtdistanz möglichst gering bleibt. In Abbildung 2.1 b) handelt es sich um das Sequenzpaar 1 und 2.

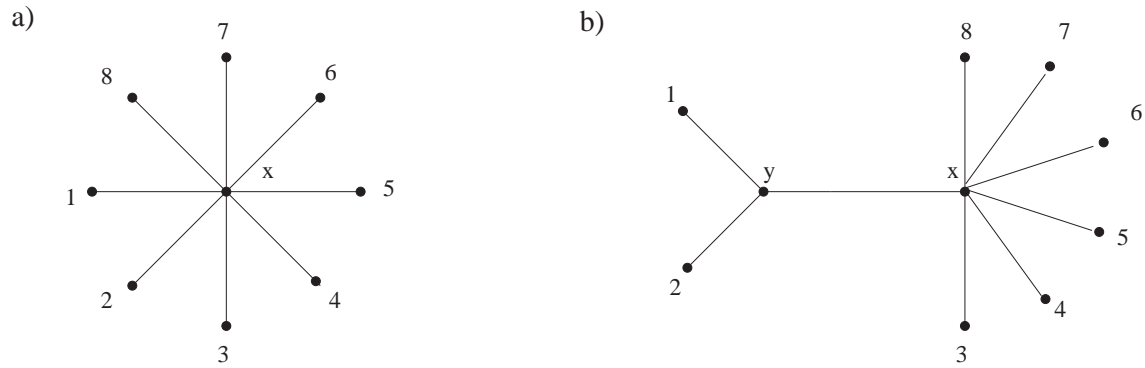


Abbildung 2.1: Der Ausgangspunkt der Neighbor Joining Methode ist eine Sterntopologie wie sie in a) zu sehen ist. Im ersten Schritt wird der innere Knoten y so eingefügt, daß durch die Wahl eines OTU Paares, in diesem Fall 1 und 2, die Gesamtdistanz des Baumes am geringsten bleibt (b).

Beim Einfügen des inneren Knotens Y muß also für jedes OTU Paar, an dem der Knoten Y plaziert werden kann, die Gesamtdistanz des Baumes berechnet werden, um die geringste Gesamtdistanz zur Positionierung von Y auszuwählen. Nach dem Einfügen muß die Distanzmatrix neu bestimmt werden. Anstatt der Distanzen zu den beiden Sequenzen des ausgewählten Paares, enthält die Matrix nun die Distanzen aller nicht ausgewählten OTUs zum inneren Knoten Y.

Für das obige Beispiel würden sich die neuen Distanzen zu dem inneren Knoten Y D_{Yk} berechnen durch (vgl. Nei and Saitu (1987)):

$$D_{Yk} = \frac{D_{1k} + D_{2k}}{2}, k \in \{3, 4, 5, 6, 7, 8\}$$

Die Schritte einen neuen inneren Knoten einzufügen und die Distanzmatrix neu zu bestimmen werden nun so lange wiederholt, bis ein phylogenetischer Baum entstanden ist. Bei jedem Einfügevorgang verringert sich die Anzahl der zu untersuchenden Knoten um 1. Der Baum ist dann vollständig erstellt, wenn die Anzahl der zu untersuchenden Knoten gleich 3 ist.

2.1.2 Maximum Likelihood Analyse mit dem Programm PROTML

2.1.2.1 Das Maximum Likelihood Prinzip

Der Likelihood eines Baumes setzt sich aus Übergangswahrscheinlichkeiten zusammen, die besagen, wie wahrscheinlich es ist, daß ein Nukleotid bzw.

eine Aminosäure durch ein anderes Nukleotid bzw. eine andere Aminosäure ersetzt wird. Hierzu werden Wahrscheinlichkeitsmatrizen der Größe 4x4 (Nukleotide) bzw. 20x20 (Aminosäuren) verwendet. Die Einträge geben jeweils die Substitutionswahrscheinlichkeiten der Nukleotide/Aminosäuren wieder.

Das folgende Beispiel, in einer leicht variierten Form aus Nei (1987) entnommen, soll verdeutlichen, wie sich mit Hilfe von Wahrscheinlichkeitsmatrizen der Likelihood eines Baumes bestimmen läßt.

Der Beispielbaum in Abbildung 2.2 sei durch ein beliebiges Verfahren für einem Alignment mit 5 Sequenzen der Länge m erstellt worden.

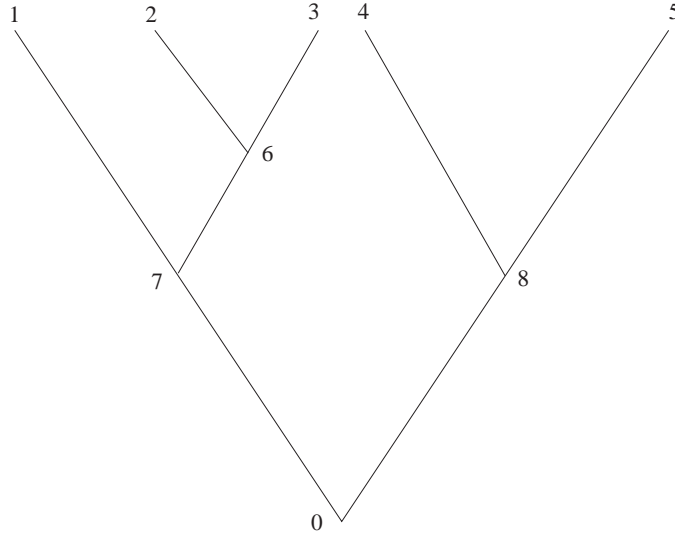


Abbildung 2.2: ML Beispielbaum mit 5 Blättern

Man betrachte eine beliebige Position i des Alignments. Unter der Annahme, daß die Sequenzen an den inneren Knoten 6, 7 und 8 bekannt sind, würde sich die Übergangswahrscheinlichkeit des Baumes an Position i L_i durch die Formel 2.1 berechnen.

$$L_i = g_{s_0[i]} P_{s_0[i], s_7[i]} P_{s_7[i], s_1[i]} P_{s_7[i], s_6[i]} P_{s_6[i], s_2[i]} P_{s_6[i], s_3[i]} P_{s_0[i], s_8[i]} P_{s_8[i], s_4[i]} P_{s_8[i], s_5[i]} \quad (2.1)$$

Der Faktor $g_{s_0[i]}$ steht für die Wahrscheinlichkeit, daß sich die Wurzel des Baumes am Knoten 0 befindet. $P_{s_k[i], s_l[i]}$ gibt die Wahrscheinlichkeit an, daß der Nukleotid- / Aminosäurezustand der Sequenz k an der Position i durch den der Sequenz l ersetzt wird.

Die Sequenzen der inneren Knoten sind nicht bekannt. Aus diesem Grunde ist es erforderlich bei der Berechnung von L_i alle möglichen Nukleotide/Aminosäuren an der Stelle zu berücksichtigen. Somit ergibt sich für L_i die Formel 2.2. Die Konstante $const$ hat hierbei den Wert 4 bei Nukleotiden oder den Wert 20 für Aminosäuren.

$$L_i = \sum_{s_0}^{const} \sum_{s_7}^{const} \sum_{s_8}^{const} \sum_{s_6}^{const} g_{s_0[i]} P_{s_0[i], s_7[i]} P_{s_7[i], s_1[i]} P_{s_7[i], s_6[i]} P_{s_6[i], s_2[i]} P_{s_6[i], s_3[i]} P_{s_0[i], s_8[i]} P_{s_8[i], s_4[i]} P_{s_8[i], s_5[i]} \quad (2.2)$$

Nach Berechnen von L_i für alle Positionen des Alignments, ergibt sich der gesamte Likelihood des Baumes aus der Summer aller L_i , wie die folgende Formel wiedergibt.

$$L = \sum_{i=1}^m L_i \quad (2.3)$$

2.1.2.2 Rekonstruktion eines ML Baums mit PROTML

Das Programm PROTML, daß im MOLPHY Paket Adachi and Hasegawa (1996) enthalten ist, bietet mehrere Optionen zur ML Analyse von Proteinalignments. Neben dem Vergleich mehrerer Topologien durch Berechnung des Likelihoods, gibt es auch Möglichkeiten zur Rekonstruktion phylogenetischer Bäume.

So geht die Option *Star – Decomposition* nach dem selben Prinzip wie die NJ Methode vor. Beginnend bei einer Sterntopologie werden die inneren Knoten immer so eingefügt, daß die entstehende Topologie den größten Likelihood aufweist.

Eine weitere Möglichkeit ist es den ML Baum ausgehend von einer vorgegeben Topologie zu bestimmen. Hierbei wird das *Local Rearrangement* angewendet (vgl. Abbildung 2.3).

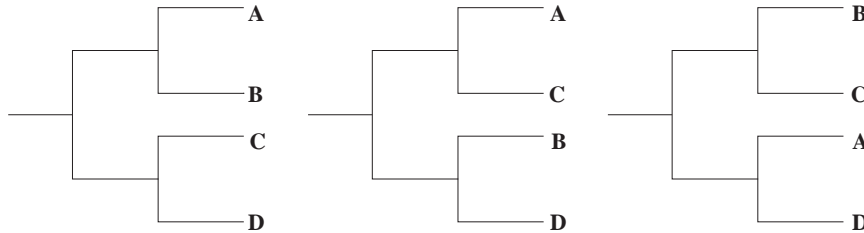


Abbildung 2.3: Das Prinzip von *Local Rearrangement*

Wie in der Abbildung 2.3 zu sehen ist, existieren an jedem inneren Knoten der

nicht ein direkter Vorgänger von zwei Blättern ist, drei mögliche Teilbäume. Beim *Local Rearrangement* werden beginnend bei den Blättern an jedem solcher inneren Knoten die Likelihood Werte für jede der drei Möglichkeiten bestimmt. Mit dem besten Teilbaum, dem mit dem höchsten Likelihood, wird an den nächst tieferen inneren Knoten das *Local Rearrangement* durchgeführt. Das Vorgehen wird so lange wiederholt, bis ein vollständiger Baum entstanden ist.

Für die ML Bäume in den Untersuchungen, die im folgendem Kapitel beschrieben werden, wurde zunächst eine NJ Topologie mit dem Programm NJDIST des MOLPHY Paketes erstellt. Ausgehend von dieser Topologie konstruierte PROTML mit dem *Local Rearrangement* Vorgehen einen ML Baum.

2.1.3 PROTDET - Vom Spektrum aller Teilbäume zum Abstammungsbaum

2.1.3.1 Definitionen

Alle Teilbäume eines Baumes, sogenannte Splits, stellen Bipartitionen dar. Ein Split ist die Menge der Blätter, die in einem Teilbaum enthalten sind. Durch einen Split ist eine Zweiteilung, eine Bipartition, des Baumes in zwei Teilbäume definiert. Dies ist zum einen der Split und zum anderen der Teilbaum, der alle übrigen Blätter des Baumes enthält. Die Bipartition besteht also aus zwei Mengen die jeweils die Blätter der Teilbäume umfassen.

Ein Baum mit N Blättern setzt sich aus $N - 3$ solcher Splits zusammen. Allerdings existieren für N Blätter 2^{N-1} mögliche Splits. Dieses Spektrum aller möglichen Splits läßt sich als Vektor mit 2^{N-1} Einträgen darstellen.

Als Beispiel sei der Baum mit 5 Blättern in Abbildung 2.4 gegeben.

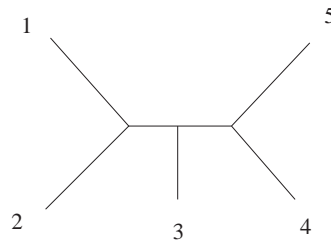


Abbildung 2.4: PROTDET Beispielbaum mit 5 Blättern

Dieser Baum ist durch die zwei Splits (1,2) und (3) definiert. Die Liste aller möglichen Splits und der dazugehörigen Bipartitionen gibt die Tabelle 2.1 wieder.

#	Split	Bipartition	Gerade Geordnete Teilmenge
0	$\{\}$	$\{\{\}, \{1,2,3,4,5\}\}$	$\{\}$
1	$\{1\}$	$\{\{1\}, \{2,3,4,5\}\}$	$\{1,5\}$
2	$\{2\}$	$\{\{2\}, \{1,3,4,5\}\}$	$\{2,5\}$
3	$\{1,2\}$	$\{\{1,2\}, \{3,4,5\}\}$	$\{1,2\}$
4	$\{3\}$	$\{\{3\}, \{1,2,4,5\}\}$	$\{3,5\}$
5	$\{1,3\}$	$\{\{1,3\}, \{2,4,5\}\}$	$\{1,3\}$
6	$\{2,3\}$	$\{\{2,3\}, \{1,4,5\}\}$	$\{2,3\}$
7	$\{1,2,3\}$	$\{\{1,2,3\}, \{4,5\}\}$	$\{1,2,3,5\}$
8	$\{4\}$	$\{\{4\}, \{1,2,3,5\}\}$	$\{4,5\}$
9	$\{1,4\}$	$\{\{1,4\}, \{2,3,5\}\}$	$\{1,4\}$
10	$\{2,4\}$	$\{\{2,4\}, \{1,3,5\}\}$	$\{2,4\}$
11	$\{1,2,4\}$	$\{\{1,2,4\}, \{3,5\}\}$	$\{1,2,4,5\}$
12	$\{3,4\}$	$\{\{3,4\}, \{1,2,5\}\}$	$\{3,4\}$
13	$\{1,3,4\}$	$\{\{1,3,4\}, \{2,5\}\}$	$\{1,3,4,5\}$
14	$\{2,3,4\}$	$\{\{2,3,4\}, \{1,5\}\}$	$\{2,3,4,5\}$
15	$\{1,2,3,4\}$	$\{\{1,2,3,4\}, \{5\}\}$	$\{1,2,3,4\}$

Tabelle 2.1: Die Tabelle enthält die für die PROTDET Analyse relevanten Mengen bezüglich eines 5 blättrigen Abstammungsbaum. In der zweiten Spalte sind alle möglichen Splitmengen aufgeführt, die für einen 5 blättrigen Baum existieren. Die Spalten drei und vier geben die zu den Splits gehörigen Bipartitionsmengen und die geordneten Teilmengen (vgl. Formel 2.4) wieder.

In der Spektralanalyse werden zu den N Blättern ein Spektrum geordneter Teilmengen definiert. Dieses Spektrum umfaßt alle Mengen, die eine gerade Anzahl von Elementen, den Blättern, enthält. Diese Mengen lassen sich leicht von den Splits ableiten. Für einen Split S_i erhält man die gerade geordnete Teilmenge E_i nach der folgenden Formel (vgl. Hendy and Penny (1993), Penny and Hendy (2001)):

$$E_i = \begin{cases} S_i \cup \{N\}, & \text{wenn } S_i \text{ eine gerade Anzahl an Elementen hat} \\ S_i, & \text{sonst} \end{cases} \quad (2.4)$$

2.1.3.2 Das PROTDET Programm

Der Versuch aus dem Spektrum von den möglichen 2^{N-1} Splits anhand geeigneter Kriterien die $N - 3$ besten, sich nicht ausschließenden Splits zu bestimmen ist die Grundidee der Spektral Analyse (Hendy and Penny (1993), Penny and Hendy (2001)), die die Grundlage des PROTDET Programmes (bisher noch nicht veröffentlicht) bildet.

Zur Gewichtung der Splits berechnet PROTDET eine Distanzmatrix, die die durch die sogenannte *logdet* Methode bestimmte Distanz zwischen allen möglichen Sequenzpaaren enthält (Steel (1994), Lockhart *et al.* (1994)). Zunächst wird aus dem Vergleich zweier Sequenzen i, j die 20x20 Matrix $\hat{D}_{ij} =$ bestimmt. Jedes Element dieser Matrix stellt den Anteil der Positionen im

gegeben Datensatz dar, an denen eine bestimmte Aminosäurekombinationen des Sequenzpaares auftritt (vgl. Formel 2.2). Folglich ergibt die Summe aller Matrixelemente, deren Werte zwischen 0 und 1 liegen, immer den Wert 1. Berechnet man nun zu dieser Matrix die Determinante und bestimmt daraus den Logarithmus ergibt dies die Distanz d_{ij} (vgl. Formel 2.3).

$$\hat{D}_{ij} = \begin{bmatrix} aa & ac & ad & \dots & \dots & av & aw & ay \\ ca & cc & cd & \dots & \dots & cv & cw & cy \\ da & dc & dd & \dots & \dots & dv & dw & dy \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ va & vc & vd & \dots & \dots & vv & vw & vy \\ wa & wc & wd & \dots & \dots & wv & ww & wy \\ ya & yc & yd & \dots & \dots & yv & yw & yy \end{bmatrix} \quad (2.5)$$

$$d_{ij} = -\ln(\det(\hat{D}_{ij})) \quad (2.6)$$

Je größer bei den verglichenen Sequenzpaar der Anteil der identischen Aminosäuren, repräsentiert durch die Hauptdiagonal aa bis yy , ist, desto größer ist auch der Wert der Determinante und umso kleiner ist der Distanzwert d_{ij} . Umgekehrt gilt, je niedriger die Werte der Hauptdiagonalen sind desto kleiner der Wert der Determinante und umso größer die Distanz.

Falls sich aus der Matrix \hat{D}_{ij} eine Determinante ergibt deren Wert kleiner oder gleich 0 ist, so läßt sich die Distanz nicht über den natürlichen Logarithmus berechnen. Bei solchen Sequenzpaaren geht das *logdet* Verfahren davon aus, daß diese Sequenzen sehr unähnlich sind und die Distanz sehr groß sein muß. Diese Paare werden absteigend nach dem Wert der Determinante sortiert, um ihnen danach einen Distanzwert zuzuweisen, der dem bisher größten Distanzwert + 0,05 entspricht.

Bei der Berechnung von \hat{D}_{ij} bietet PROTDET die Möglichkeit den Positionen des Alignments ein Gewicht zuzuweisen. In der Grundeinstellung werden alle Positionen gleich, mit den Wert 1 gewichtet. Optional lassen sich auch konstante, variable oder bestimmten Splits entsprechende Positionen mit Werten kleiner 1 gewichten. Die für die Untersuchungen im Kapitel 5 wurde die *downweight* Option angewendet. Hierbei werden 11 PROTDET Berechnungen durchgeführt, wobei die Gewichtungen konstanter Positionen jeweils in gleichgroßen Schritten reduziert wird. Das Anfangs- und das Endgewicht kann dabei vom Benutzer vorgegeben werden. Das „Heruntergewichten“ bewirkt, daß die Berücksichtigung der konstanten Positionen bei der Analyse schrittweise verringert wird.

Zu den 2^{N-1} Split werden 2^{N-1} Teilmengen definiert, die jeweils eine gerade Anzahl an Taxa enthält (siehe oben). Aus der Distanzmatrix läßt sich der Distanzvektor bezüglich dieser Teilmengen berechnen. Dieser enthält für jeden der 2^{N-1} Teilmengen die minimale Summe aller paarweise Distanzen der

Teilmenge, die aus der Matrix resultieren. Aus diesem Vektor soll nun der Vektor erstellt werden, der für jeden Split seine Frequenz bzw. den Anteil der Positionen im Datensatz enthält, die dem Split im Bezug auf Distanzmatrix wiedergibt.

Um vom Distanzvektor zum Frequenzvektor zu gelangen, wird eine Hadamard Transformation durchgeführt (Penny and Hendy (2001), Hendy and Penny (1993)). Hierzu wird der Distanzvektor mit der Hadamard-Matrix multipliziert. Diese quadratische Matrix hat entsprechend dem Vektor die Größe $2^{N-1} \times 2^{N-1}$. Jedes Element der Matrix hat den Wert 1 oder -1. Alle Zeilen sowie alle Spalten sind orthogonal zueinander, d.h. daß die Summe aller paarweisen Produkte korrespondierender Zeilen-/Spaltenelemente den Wert 0 ergibt.

Durch die Multiplikation des Vektors mit der Hadamardmatrix werden die Distanzen so addiert bzw. subtrahiert, daß sich für jeden der 2^{N-1} Splits die entsprechende Frequenz ergibt.

Je höher die Frequenz für einen Split desto besser spiegelt die Distanzmatrix diesen Split wider. Allerdings ist es auch möglich, daß eine negative Frequenz berechnet wird. Das bedeutet, daß aus der Distanzmatrix Informationen hervorgehen, die dem Split stark widersprechenden.

Aus diesem Spektrum lassen sich mit unterschiedlichen Kriterien („closest tree selection“, parsimony, ...) die „besten“ Splits auswählen, aus denen sich der resultierende Baum zusammensetzt (Hendy and Penny (1993)).

Bei den Untersuchungen in Kapitel 5 wurde die „closest tree selection“ verwendet. Hierbei werden aus dem Baum Spektrum die $N - 3$ sich nicht ausschließenden Splits mit der insgesamt höchsten Wahrscheinlichkeit ausgewählt.

2.1.4 RASA - Relative Apparent Synapmophy Analysis

Die in diesem Abschnitt beschriebene RASA Analyse (Lyons-Weiler and Hoelzer (1996), Lyons-Weiler and Hoelzer (1997), Lyons-Weiler and Hoelzer (1999)) versucht eine Entscheidung darüber zu treffen, ob ein Datensatz ein signifikantes phylogenetisches Signal enthält, ohne dabei eine Aussage über eine mögliche Baumtopologie zu erstellen.

Das grundlegende Kriterium für das Vorhandensein eines phylogenetischen Signals besteht bei RASA darin, inwieweit die Daten die Existenz von Sequenztriplets der Form $((i,j),k)$ für Sequenzen i,j,k mit $i \neq j \neq k$ beinhalten.

Hierbei wird zunächst für jedes Sequenzpaar i,j , mit $i \neq j$ an jeder Sequenzposition p des Alignments überprüft, ob die Sequenzen übereinstimmen, also $S_i[p] = S_j[p]$ gilt. Die Anzahl der identischen Positionen gibt für jedes Sequenzpaar der Wert E_{ij} wieder. An jeder Position an dem eine solche

Übereinstimmung vorliegt wird weiterhin die Anzahl der übrigen Sequenzen bestimmt, die sich vom Sequenzpaar (i, j) unterscheiden. Der Wert $Rasa_{ij}$ beinhaltet die Summe dieser Sequenzen bezüglich eines Paares (i, j) im gesamten Datensatz.

Stellt man nun für alle Sequenzpaare die Werte $Rasa_{ij}$ (Anzahl der Triplets) über den entsprechenden E_{ij} (Anzahl der Paare) in einem Koordinatensystem dar, läßt sich eine lineare Abhängigkeit erkennen, für die mit Hilfe der linearen Regression eine Ausgleichgerade bestimmt werden kann (vgl. Bronstein and Semendjajew (1991), Lozan and Kausch (1998)). Diese hat die folgende Form einer linearen Geradengleichung.

$$y = \beta * x + \alpha \quad (2.7)$$

Um eine Aussage geben zu können, ob diese Gerade ein signifikantes Signal darstellt, wird eine Vergleichsgerade, eine sogenannte Nullgerade, verwendet, die durch den Ursprung des Achsenkreuzes verläuft. Diese Gerade soll den Fall repräsentieren, daß das gegebene Alignment kein phylogenetisches Signal enthält. Um diese zu erhalten wird in Lyons-Weiler and Hoelzer (1999) vorgeschlagen, bezüglich des gegebenen Alignments eine Anzahl von Alignments mit gleicher Sequenz- und Positionsanzahl zu erstellen. Jede der Positionen in diesen Datensätzen entspricht einer Permutation der entsprechenden Position aus dem gegebenen Daten. Mit dem oben beschriebenen Vorgehen läßt sich für jedes dieser Alignments durch den Sequenzvergleich die Werte E_{ij} und $Rasa_{ij}$ bestimmen, sowie die dazugehörige Ausgleichsgerade berechnen. Der Steigungsfaktor β_0 der Nullgerade wird nun durch den Mittelwert aller Steigungsfaktoren β gebildet. Die Nullgerade hat damit die Form

$$y = \beta_0 * x \quad (2.8)$$

Mit dem statistischen t-Test (vgl. Bronstein and Semendjajew (1991), Lozan and Kausch (1998)) können nun die beiden Geradengleichungen verglichen werden. Hierzu benötigt man zwei Werte. Zum einen t_{rasa} dessen Berechnung in den Gleichungen 3.5 bis 3.7 wiedergegeben ist (siehe auch Lozan and Kausch (1998)). Hierbei bedeuten die Variablen \bar{x} und \bar{y} die Mittelwerte aller x (E_{ij}) bzw. y (T_{Rasa}) Werte und n entspricht der Sequenzanzahl des Alignments. Zum anderen den kritischen Schwellwert t_{crit} . Dieser ergibt sich aus der Studentschen Verteilung mit dem Freiheitsgrad $((n * n - n)/2) - n - 3$ (Lyons-Weiler and Hoelzer (1996), Lyons-Weiler and Hoelzer (1997)). Ist $t_{rasa} > t_{crit}$, so enthält der untersuchte Datensatz ein signifikantes phylogenetisches Signal.

$$t_{rasa} = \frac{\beta - \beta_0}{S_\beta} \quad (2.9)$$

$$S_\beta = \frac{S_{xy}}{\sqrt{\sum (x_i - \bar{x})^2}} \quad (2.10)$$

$$S_{xy} = \frac{\sum (y_i - \bar{y})^2 - \frac{(\sum x_i y_i - \frac{\sum x_i \sum y_i}{n})^2}{\sum (x_i - \bar{x})^2}}{n - 2} \quad (2.11)$$

Der RASA Algorithmus war bisher nur mit einem Programm für Macintosh Rechner verfügbar. Unter <http://bio.uml.edu/LW/RASA.html> sind Verweise zum Download der aktuellen Version und einem Online Tool Server zu finden.

Um RASA auch für das Betriebssystem LINUX nutzbar zu machen, wurde RASA für diese Dissertation in einer leicht variierten Version neu implementiert. Beschreibungen hierzu sind im Anhang A und B wiederzufinden.

2.1.5 Die Variabilität von Sequenzpositionen als Maß für phylogenetische Information

Ein bekanntes Problem bei der Rekonstruktion phylogenetischer Bäume ist es, daß die Alignments phylogenetisch informative aber auch wenig informative Daten enthalten, die teilweise auch einen negativen Einfluß auf die Rekonstruktion der Abstammungsbäume haben. Die in diesem Abschnitt vorgestellten Vorgehensweisen stellen einen „intuitiven“ Ansatz vor, die versuchen, die Daten eines Alignments in Klassen einzuteilen, um die „störenden“ Daten lokalisieren und von weiteren phylogenetischen Untersuchungen ausschließen zu können.

2.1.5.1 Variabilitätsklassen eines Datensatzes

In einem Alignment lassen sich viel oder kaum veränderte Sequenzabschnitte anhand der jeweiligen Anzahl unterschiedlicher Aminosäuren erkennen. Diese Anzahl wird als Variabilität einer Position bezeichnet. Die Variabilität nimmt den Wert 1 für konservierte Positionen an, bei denen die Aminosäuren in allen Sequenzen übereinstimmen. Mit zunehmender Zahl unterschiedlicher Aminosäuren erhöht sich auch der Wert der Variabilität. Falls das Alignment weniger als 20 Sequenzen umfaßt, so kann die Variabilität maximal den Wert der Sequenzanzahl annehmen, bei mehr als 20 Sequenzen den liegt das Maximum bei 20, der Anzahl der verschiedenen Aminosäuren. Mit einer Einteilung in Variabilitätsklassen lassen sich stark bzw. wenig veränderte Sequenzabschnitte zusammenfassen. Die Positionen der Variabilitätsgruppe 1 repräsentieren alle konservierten, also nicht veränderten Sequenzabschnitte, die Variabilitätsklasse mit dem maximalen Variabilitätswert umfaßt die am meisten veränderten Positionen.

Geht man nun von der Annahme aus, daß wenig veränderte Abschnitte tendenziell zu wenig phylogenetische Informationen liefern und sehr häufig veränderte Positionen eher einen störenden Einfluß auf die Rekonstruktion der Abstammungsbäume haben, bietet die Einteilung in Variabilitätsklassen einen geeigneten Ausgangspunkt phylogenetisch relevante Daten in einem Alignment

zu lokalisieren.

2.1.5.2 Das Programm SORTAL

Im Rahmen der Arbeiten Hansmann (1997) und Hansmann (1998) wurde SORTAL selbst entwickelt. Das Programm SORTAL - Sort Alignment - nimmt die Einteilung und Sortierung eines Alignments nach den oben beschriebenen Variabilitätsklassen vor. Mit SORTAL werden die Positionen eines Alignments im PHYLIP Format (Felsenstein (1993)) durch die Variabilität klassifiziert und in einem PHYLIP Alignment in sortierter Reihenfolge beginnend bei den konserviertesten bis hin zu den variabelsten Positionen ausgegeben. Bei der Klassifizierung werden “gaps“ als variabel, also wie zu den anderen Aminosäuren verschieden gehandhabt. Die höchste Variabilitätsklasse kann aber auch dann nur maximal den Wert 20 annehmen, wenn das Alignment mehr als 20 Sequenzen beinhaltet.

Um Positionen gleicher Variabilität ebenfalls sortieren zu können, wird an den Positionen jeweils die Anzahl der dort am seltensten, aber mindestens einmal auftretenden Aminosäure *minaa* als Kriterium gewählt. Im sortierten Alignment werden die Positionen gleicher Variabilität beginnend mit dem kleinsten *minaa* Wert ausgegeben.

Realisiert wird die Sortierung dadurch, daß zu jeder Position eine Priorität wie in folgender Gleichung bestimmt wird.

$$priority = 100 * variability + minaa, minaa > 0$$

Anhand der Priorität werden die Positionen aufsteigend in eine Liste einsortiert und in einem Alignment ausgegeben.

Das folgende Beispiel soll die Klassifizierung und Sortierung eines Alignments verdeutlichen. Der Beispieldatensatz umfaßt 17 Sequenzen jeweils der Länge von 48 Aminosäuren (s. Abbildung 2.5).

In Abbildung 2.6 ist das sortierte Alignment zu sehen. Es lassen sich die folgenden Variabilitätsklassen daraus erkennen.

Variabilität	Positionen
1	1 bis 8
2	9 bis 20
3	21 bis 28
4	29 bis 31
5	32 bis 34
6	35 bis 37
17	38 bis 48

17	48	
seqA	MKIRASVKKI	CNQ---CRLI RRGKILIIC K-NPKHKQRQ IKKIKT--
seqB	MKVRASVRKI	CSR---CVAL KRHGV LIVIC S-NPKHKQRQ G-----
seqC	MKVRPSVKKM	CDK---CRLI RRGKTLRVIC Q-NPKHKQRQ G-----
seqD	MKVRASVRKM	CEK---CRTI RRGGRVMVIC S-NSKHKQRQ G-----
seqE	MKIRSSVKKI	CNK---CYLI RRKNLLVVC I-NNKHKQRQ G-----
seqF	MKVSSIGSL	KNRSKDCQIV KRRGRIYVIC KTDPR LKVRQ GGAKMKRK
seqG	MKIRASVRKI	CTK---CRLI RRRGRIRVIC S-NPKHKQRQ G-----
seqH	MKIRASVRKI	CEN---CRLI RRRRIMVVC S-NPKHKQRQ G-----
seqI	MKVRASVRKI	CEN---CRLI KRRGTVMVIC SNNPKHKQRQ G-----
seqJ	MKVRPSVRKI	CDK---CCLI RRHRKLLVIC S-NPKHKQRQ G-----
seqK	MKIRASVRKI	CEK---CRLI RRRGRIIVIC S-NPRHKQRQ G-----
seqL	MKVRPSVKKM	CDK---CRVI RRGKIMVIC P-NAKHKQRQ G-----
seqM	MKIRASVRKI	CTK---CRLI RRRGRIIVIC S-NPRHKQRQ G-----
seqN	MKIRASVRKI	CTK---CRLI RRRGRIRVIC S-NPKHKQRQ G-----
seqO	MKIRASIRRI	CGK---CRPI RRRKRVMIIC S-NPRHKQKQ G-----
seqP	MKVRPSVRKM	CEK---CRII RRHRKVMVIC N-NPKHKQRQ G-----
seqQ	MKVRASVKKM	CDK---CRVI RRRGRVMVIC SANPKHKQRQ G-----

Abbildung 2.5: Beispielalignment vor dem Sortieren

2.1.5.3 Phylogenetische Rekonstruktion einzelner Variabilitätsklassen mit PBL

Das Programm PBL, wie SORTAL selbst implementiert (vgl. Hansmann (1997), Hansmann (1998)), bietet die Möglichkeit aus einem mit SORTAL sortierten Alignment variabilitätsabhängige Teilalignments systematisch mit drei unterschiedlichen phylogenetischen Programmen zu untersuchen.

Bei den drei Methoden handelt es sich neben den in Abschnitt 2.1.2 vorgestellten PROTML Adachi and Hasegawa (1996) Vorgehen um die PHYLIP Programme (Felsenstein (1993)) PROTPARS und NJ-Methode. In Hansmann (1998) wurden diese Programme bereits quantitativ und qualitativ verglichen. Für die Analysen im folgenden Kapitel wurde PROTML verwendet. Es liefert zum einen gute und verlässliche Ergebnisse und ist zudem noch weitaus weniger rechenintensiv als die beiden PHYLIP Programme.

Mit dem Programm lassen sich auf 4 verschiedene Arten Teilabschnitte aus dem sortierten Datensatz zur Rekonstruktion heranziehen. In Hansmann (1997) und Hansmann (1998) wurden bereits zwei davon vorgestellt.

Die eine Möglichkeit verkürzt den Datensatz um eine vom Benutzer definierte konstante Anzahl der variabelsten Positionen. Bei der anderen Möglichkeit wurde jeweils ein vom Benutzer konstant definierter Teilbereich des Alignments von *lw* Positionen bei der phylogenetischen Rekonstruktion in Betracht gezogen. Dieser Teilbereich, ein sogenanntes Fenster, wird zunächst aus den variabelsten Positionen erstellt, bevor das Fenster in Richtung der konserviertesten Positionen um *shift* Positionen verschoben wird. Der *shift* Faktor wird ebenfalls vom Anwender definiert.

2.2 Die untersuchten Datensätze

In diesem Abschnitt sollen die Datensätze vorgestellt werden, die als Grundlage der Untersuchungen in dieser Arbeit dienen.

Der Gegenstand der Analysen waren zwei Datensätze. Bei beiden Datensätzen handelt es sich um Alignments, die aus der Konkatenierung einzelner Proteinalignments zusammengestellt wurden. Der Bakteriendatensatz beinhaltet 32 homologe Gene aus 26 Bakteriengenomen und der Plastidendatensatz 41 Plastidengene aus 16 Organismen. Zur Erstellung der Alignments wurden jeweils die Aminosäuresequenzen der homologen Gene verwendet.

2.2.1 Der Bakteriendatensatz

Die Ausgangspunkt zur Erstellung dieser Alignments war ein Artikel von Wächtershäuser (Wächtershäuser (1998)), der ein Cluster in der Anordnung von etwa 50 Genen in 19 zum größten Teil vollständig sequenzierten Bakteriengenomen untersucht hat.

Zunächst wurden aus den bereits vollständig sequenzierten Genomen des Artikels die Gene identifiziert, die sowohl bei den Archae- wie auch bei den Eubakterien Homologe aufweisen. In Abbildung 2.7 sind die konserviertesten Abschnitte des Wächtershäuser Clusters dargestellt. Die zwei *Sulfolobus* Arten wurden dabei nicht berücksichtigt, weil die Genome noch nicht vollständig sequenziert sind. In der Tabelle sind Proteine, die innerhalb oder außerhalb des Clusters bzw. gar nicht im Genom vorhanden sind, durch unterschiedliche Farben hervorgehoben. Die aus dem Cluster ausgewählten Proteine sind durch rote Pfeile markiert.

Die 17 Genome des Clusters wurden aus Datenbanken heruntergeladen. Mit Hilfe von BLAST Suchprogrammen (GCG (1994)) ließen sich die entsprechenden Proteine lokalisieren und als einzelne Sequenzdateien abspeichern.

Aus diesen Sequenzen wurden mit den Programmen *pileup* (GCG (1994)) oder *CLUSTALW* (Thompson *et al.* (1994)) für jedes Gen ein Proteinalignment erstellt und mit PROTML (Adachi and Hasegawa (1996)) jeweils ein phylogenetischer Baum berechnet. Aus den Alignments, deren Bäume eine Trennung zwischen Archaeobakterien und Eubakterien aufwiesen, entstand durch Aneinanderfügen ein Gesamtalignment.

Mit diesem Vorgehen wurden die Alignments jeweils um aktuell veröffentlichte Bakteriengenome ergänzt und zu einem Gesamtalignment konkateniert. Die Alignments boten jeweils Grundlage für phylogenetische Untersuchungen (Martin and Hansmann (2000)) wie sie im Kapitel 3 vorgestellt werden.

Zusätzlich zu den 32 Proteinalignments und dem daraus resultierenden Gesamtalignment aus den in Tabelle 2.2 wurden entsprechende Datensätze

erstellt, die nur die Eubakterien berücksichtigten. Eine Übersicht zu den erstellten Alignments ist in Tabelle 2.3 dargestellt.

Eubakterien	Abkürzungen	Referenz	kbp	Anzahl Gene
<i>Aquifex aeolicus</i>	aae	Deckert <i>et al.</i> (1998)	1551,335	1522
<i>Borrelia burgdorferi</i>	bb	Fraser <i>et al.</i> (1997)	910,724	850
<i>Bacillus subtilis</i>	bs	Kunst <i>et al.</i> (1997)	4214,814	4100
<i>Campylobacter jejuni</i>	cj	Parkhill <i>et al.</i> (2000)	1641,481	1654
<i>Chlamydia pneumoniae</i>	cp	Kalman <i>et al.</i> (1999)	1230,230	1052
<i>Chlamydia trachomatis</i>	ct	Stephens <i>et al.</i> (1998)	1042,519	894
<i>Deinococcus radiodurans</i> *)	der	White <i>et al.</i> (1999)	2648,638	2580
			+412,348	+357
<i>Escherichia coli</i>	ec	Blattner <i>et al.</i> (1997)	4639,221	4289
<i>Haemophilus influenzae</i>	hi	Fleischmann <i>et al.</i> (1995)	1830,138	1709
<i>Helicobacter pylori</i>	hp	Tomb <i>et al.</i> (1997)	1667,867	1566
<i>Mycoplasma genitalium</i>	mg	Fraser <i>et al.</i> (1995)	580,074	480
<i>Mycoplasma pneumoniae</i>	mp	Himmelreich <i>et al.</i> (1996)	816,394	688
<i>Mycobacterium tuberculosis</i>	mtu	Cole <i>et al.</i> (1998)	4411,529	3918
<i>Neisseria meningitidis</i>	nm	Tettelin <i>et al.</i> (2000)	2272,351	2025
<i>Pseudomonas aeruginosa</i>	pae	Stover <i>et al.</i> (2000)	6264,403	5565
<i>Rickettsia prowazekii</i>	rp	Andersson <i>et al.</i> (1998)	1111,523	834
<i>Synechocystis sp.</i>	ssp	Kaneko <i>et al.</i> (1996)	3573,470	3169
<i>Thermotoga maritima</i>	tm	Nelson <i>et al.</i> (1999)	1860,725	1846
<i>Treponema pallidum</i>	tp	Fraser <i>et al.</i> (1998)	1138,011	1031
<i>Vibrio cholerae</i> *)	vib	Heidelberg <i>et al.</i> (2000)	2961,149	2736
			+1072,315	+1092
<i>Xylella fastidiosa</i>	xyf	Simpson <i>et al.</i> (2000)	2679,306	2766
Archaeobakterien	Abkürzungen	Referenz	kbp	Anzahl Gene
<i>Aeropyrum pernix</i>	aep	Kawarabayasi <i>et al.</i> (1999)	1669,695	2694
<i>Archaeoglobus fulgidus</i>	af	Klenk <i>et al.</i> (1997)	2178,400	2420
<i>Methanococcus jannaschii</i>	mj	Bult <i>et al.</i> (1996)	1664,970	1715
<i>Methanobacterium thermoautotrophicum</i>	mtu	Smith <i>et al.</i> (1997)	1751,377	1869
<i>Pyrococcus horikoshii</i>	ph	Kawarabayasi <i>et al.</i> (1998)	1738,505	2064

Tabelle 2.2: Übersicht der 26 Bakterien bezüglich ihrer Klassifizierung, der verwendeten Abkürzungen und der Referenz für die im Datensatz enthaltenen Gene. Unter den Eubakterien sind die Gruppierungen **Gram-positiv**, **Mycoplasma-Gruppe**, **Cyanobakterien**, **obligat parasitische Bakterien** und **Spirochaeten** farblich hervorgehoben. Bei den mit *) markierten Genomen handelt es sich um Chromosom 1 und 2.

Proteinalignments	Anzahl der Positionen in den Alignments mit Archae- und Eubakterien	Anzahl der Positionen in den Alignments mit Eubakterien
rpl1	259	252
rpl2	317	312
rpl3	430	302
rpl4	280	258
rpl5	232	202
rpl6	231	231
rpl7/12	140	138
rpl10	372	215
rpl11	193	164
rpl13	205	185
rpl14	157	137
rpl15	205	177
rpl17	210	213
rpl18	217	130
rpl22	249	218
rpl24	151	147
rps3	320	304
rps5	261	247
rps7	251	184
rps8	178	177
rps9	194	170
rps10	123	122
rps11	159	152
rps12	174	156
rps13	157	127
rps17	164	164
rps19	202	193
efg	794	733
eftu	563	527
nusg	526	522
rpoa	451	450
secY	617	517
Gesamtalignment	8982	8026

Tabelle 2.3: Übersicht der einzelnen Bakterien Alignments

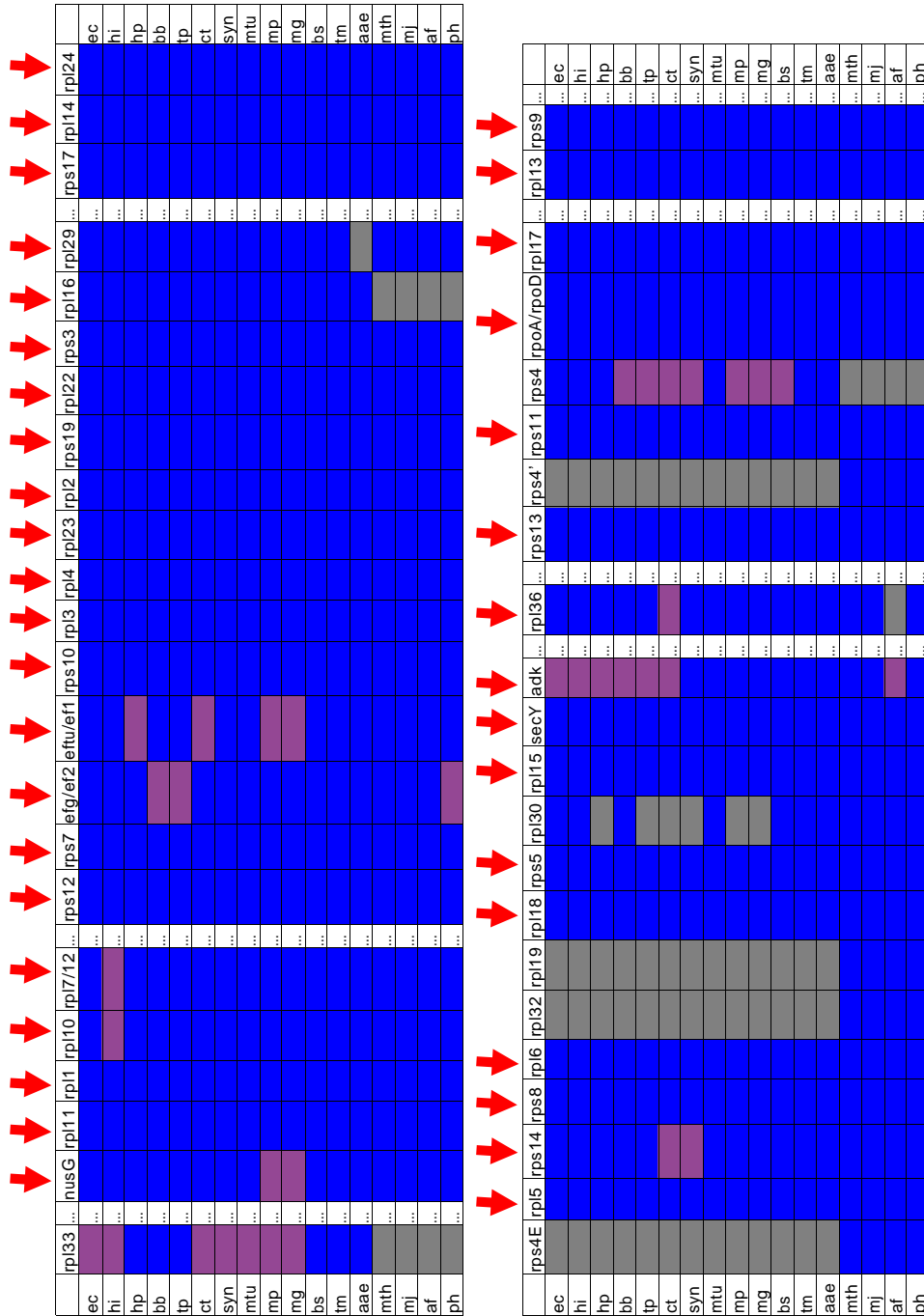


Abbildung 2.7: Die Tabelle gibt die konserviertesten Abschnitte des Wächtershäuser Clusters Wächtershäuser (1998) ohne die *Sulfolobus* Arten wieder. Die Einträge in dieser Tabelle wurden gegenüber dem Wächtershäuser Cluster in einigen Fällen korrigiert und um die fehlenden Einträge bei *Thermotoga* ergänzt. Die Einträge haben folgende Bedeutungen: blau - Gen ist im Cluster vorhanden; rot - Gen außerhalb des Clusters vorhanden; grau - Gen im Genom nicht vorhanden. Die roten Pfeile markieren die Gene, die für die Erstellung des ersten konkatenierten Datensatzes ausgewählt wurden.

	Abkürzung	Referenz	kbp	Anzahl Gene
<i>Cyanidium caldarium</i>	can	Vogel <i>et al.</i> (1996)	164,921	197
<i>Chlorella vulgaris</i>	chl	Wakasugi <i>et al.</i> (1997)	150,613	174
<i>Cyanophora paradoxa</i>	cya	Stirewalt <i>et al.</i> (1995)	135,599	149
<i>Euglena gracilis</i>	eug	Hallick <i>et al.</i> (1993)	143,172	66
<i>Guillardia theta</i>	gui	Douglas and Penny (1999)	121,524	147
<i>Zea mays</i>	mai	Maier <i>et al.</i> (1995)	140,384	111
<i>Marchantia polymorpha</i>	mar	Ohyama <i>et al.</i> (1986)	121,024	89
<i>Mesostigma viride</i>	mes	Lemieux <i>et al.</i> (2000)	118,360	105
<i>Nephroselmis olivacea</i>	nep	Turmel <i>et al.</i> (1999)	200,799	155
<i>Nicotiana tabacum</i>	nic	Shinozaki <i>et al.</i> (1986)	155,939	102
<i>Odontella sinensis</i>	odo	Kowallik <i>et al.</i> (1995)	119,704	140
<i>Oenothera elata</i>	oen	Hupfer <i>et al.</i> (2000)	159,443	123
<i>Oryza sativa</i>	ory	Hiratsuka <i>et al.</i> (1989)	134,525	108
<i>Pinus thunbergii</i>	pin	Wakasugi <i>et al.</i> (1994)	119,707	160
<i>Porphyra purpurea</i>	por	Reith and Mumholland (1995)	191,028	209
<i>Synechocystis sp.</i>	syn	Kaneko <i>et al.</i> (1996)	3573,470	3169

Tabelle 2.4: Der Plastidendatensatz enthält 41 homologe Gene von der Cyanobakterie *Synechocystis sp.*, 6 **Landpflanzen** und 10 Algenarten (darunter 2 Rotalgen, 1 Kieselalge, 2 Grünalgen)

2.2.2 Der Plastidendatensatz

Als Gegenstand zur Untersuchung der Plastidenphylogenie umfaßt diese Datenmenge 41 homologe Gene von 6 Landpflanzen, 10 Algenarten und der Cyanobakterie *Synechocystis sp.* (vgl. Tabelle 2.4).

Aus den 15 vollständig sequenzierten Chloroplastgenomen sowie den *Synechocystis*genom konnten nun die Proteine mit Hilfe der BLAST (GCG (1994)) Suche lokalisiert werden, die für alle 16 Organismen Homologe aufweisen. Eine Übersicht zu den verschiedenen identifizierten 276 Proteinen der Chloroplastgenome bezüglich der Organismen gibt die Tabelle 2.5 wieder. Das Vorhandensein eines Proteins ist durch ein + das Fehlen durch ein – markiert. Die farblich hervorgehobenen Proteine weisen Homologe in allen Genomen auf. In früheren Analysen (Martin *et al.* (1998)) wurde für jedes dieser Proteine ein Datensatz und daraus eine phylogenetischer Baum erstellt. Es zeigte sich, daß sich die Topologien der blau markierten Proteine *rpoB*, *rpoC1* und *rpoC2* deutlich von den übrigen, rot markierten Proteinen unterschieden. Aus diesem Grund wurden diese drei Proteine beim Erstellen des konkatenierten Datensatzes nicht mehr berücksichtigt.

	Z e a	O r y	N i c	O e n	P i n	M a r	E u g	C h l	N e p	M e s	O d o	G u i	P o r	C a n	C y a	S y n
accA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
accB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
accD	-	-	+	+	+	+	-	+	+	-	-	-	+	+	-	+
acpP	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
apcA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
apcB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
apcD	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
apcE	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
apcF	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
argB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
atpA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
atpB	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
atpD	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
atpE	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
atpF	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
atpG	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
atpH	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
atpI	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+
basI	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-	+
bioY	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
carA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
cbbX	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-
ccsA	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
cemA	+	+	+	+	+	+	-	+	+	+	-	+	+	+	-	+
chlB	-	-	-	-	+	+	-	+	+	+	-	-	+	-	+	+
chlI	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
chlL	-	-	-	-	+	+	-	+	+	+	-	-	+	+	+	+
chlN	-	-	-	-	+	+	-	+	+	+	-	-	+	-	+	+
clpC	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
clpP	+	+	+	+	+	+	-	+	+	+	-	-	-	-	+	+
cobA	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
cpcA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
cpcB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
cpcG	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
cpeA	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
cpeB	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	+
crtE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
cvuI	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-
cysA	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	+
cysT	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	+
desA	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
dfr	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
dnaB	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
dnaK	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
dsbD	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
fabH	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
fdx	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ftsB	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
ftsI	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+
ftsH	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
ftsW	-	-	-	-	-	-	-	-	+	+	-	-	-	-	+	+
glnB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
gltB	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
nodM	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
groEL	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
groES	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
hemA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
hisH	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
hlpA	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+

[illegible]

	Z e a	O r y	N i c	O e n	P i n	M a r	E u g	C h l	N e p	M e s	O d o	G u i	P o r	C a n	C y a	S y n
rps8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps9	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
rps10	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
rps11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps12	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps13	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
rps14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps15	+	+	+	+	+	+	-	-	-	+	-	-	-	-	-	+
rps16	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+
rps17	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+
rps18	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps19	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
rps20	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+
secA	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
secY	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
syfB	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
syh	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
thdF	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
thiG	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+
trpA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
trpG	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
trxA	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
tsf	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
tufA	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
upp	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf1	-	-	+	+	+	+	-	-	+	+	-	-	-	-	-	-
ycf2	-	-	+	+	+	+	-	-	+	+	-	-	-	-	-	-
ycf3	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
ycf4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ycf6	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+
ycf9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ycf12	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+
ycf13	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	(+)
ycf15	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf16	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
ycf17	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
ycf19	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
ycf20	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	+
ycf21	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+
ycf22	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf23	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+
ycf24	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
ycf27	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
ycf29	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
ycf32	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
ycf33	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+
ycf34	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+
ycf35	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
ycf36	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+	+
ycf37	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+
ycf38	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+
ycf39	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+
ycf40	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+
ycf41	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-	+
ycf43	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
ycf44	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
ycf45	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	+
ycf46	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
ycf47	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+

	Z e a	O r y	N i c	O e n	P i n	M a r	E u r	C h l	N e p	M e s	O d o	G u i	P o r	C a n	C y a	S y n
ycf48	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
ycf49	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
ycf50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
ycf51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
ycf52	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf53	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
ycf54	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
ycf55	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
ycf56	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf57	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf58	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
ycf59	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+
ycf60	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf61	-	-	-	-	-	-	-	-	+	-	+	+	+	+	-	+
ycf62	-	-	-	-	-	-	-	+	+	+	-	-	+	+	-	+
ycf63	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf64	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+
ycf65	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	+
ycf66	-	-	-	-	-	+	-	-	-	+	+	-	-	-	-	+
ycf67	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-
ycf68	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-
ycf69	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf70	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf71	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf72	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf73	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf74	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf75	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf76	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ycf77	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-
ycf78	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-
ycf79	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
ycf80	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
ycf81	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	+
ycf82	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
ycf83	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
ycf84	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+
rne	-	-	-	-	-	-	-	-	+	-	-	+	+	-	-	+
rnpB	-	-	-	-	-	-	-	-	+	-	-	-	+	-	+	+

Tabelle 2.5: Kodierung der Proteine in sequenzierten Chloroplastengenomen und der Cyanobakterie *Synechocystis*. Mit + wird das Vorhandensein, mit – die Abwesenheit der Proteine in den Genomen markiert. Die Proteine, die in allen Genomen Homologe ausweisen sind farblich hervorgehoben. Aus den mit rot markierten Proteinen ist das konkatenierte Alignment zusammengesetzt. Die Alignments der mit blau hervorgehobenen Proteine *rpoB*, *rpoC1* und *rpoC2* sind nicht Bestandteil des konkatenierten Datensatzes. Wie in Martin *et al.* (1998) beschrieben, ergaben phylogenetische Untersuchungen der einzelnen Proteinalignments für diese 3 Proteine Topologien, die im Widerspruch zu den übrigen Bäumen der Proteindatensätze stehen. Zea - *Zea mays*, Ory - *Oryza sativa*, Nic - *Nicotiana tabacum*, Oen - *Oenothera elata*, Pin - *Pinus thunbergii*, Mar - *Marchantia polymorpha*, Eug - *Euglena gracilis*, Chl - *Chlorella vulgaris*, Nep - *Nephroselmis olivacea*, Mes - *Mesostigma viride*, Odo - *Odontella sinensis*, Gui - *Guillardia theta*, Por - *Porphyra purpurea*, Can - *Cyanidium caldarium*, Cya - *Cyanophora paradoxa*, Syn - *Synechocystis sp.*

Proteinalignments	Anzahl der Positionen im Alignment	Proteinalignments	Anzahl der Positionen im Alignment
atpa	507	psbl	39
atpb	501	psbn	45
atpe	143	psbt	35
atpf	192	rpl2	29
atph	82	rpl14	124
petb	222	rpl16	143
petg	39	rpl20	131
psaa	755	rpl36	48
psab	738	rps2	279
psac	82	rps3	291
psaj	45	rps4	215
psba	361	rps7	159
psbb	509	rps8	148
psbc	487	rps11	146
psbd	355	rps12	12
psbe	85	rps14	103
psbf	44	rps18	175
psbh	82	rps19	96
psbi	53	ycf4	214
psbj	42	ycf9	110
psbk	61	Gesamtalignment	8308

Tabelle 2.6: Übersicht der einzelnen Alignments plastidärer Proteine

Das Plastidenalignment wurde wie im vorherigen Abschnitt beschrieben erstellt. Aufbauend auf den Datensätzen, die als Grundlage für die Analysen in Martin *et al.* (1998) sowie Stoebe *et al.* (1999) dienten, entstand dieser Datensatz durch Ergänzung aktuell veröffentlichter Plastidengenome (siehe auch Tabelle 2.4).

Die 41 Proteinalignments aus denen sich das Plastidenalignment zusammensetzt ist unter der Angabe der jeweiligen Positionenanzahl in Tabelle 2.6 dargestellt.

Kapitel 3

Ergebnisse

Wie im vorherigen Kapitel beschrieben handelt es sich bei den Datensätzen um konkatenierte Proteinalignments. Die einzelnen Proteine eines Genoms unterlagen unterschiedlichen Evolutionsprozessen. Somit lassen einzeln untersuchte Proteinalignments nur schwer Rückschlüsse auf die phylogenetischen Beziehungen zwischen den zu vergleichenden Spezies zu. Die Konkatenierung der Proteindaten zu einem einzigen Datensatz wirkt diesem Problem entgegen, weil so die verschiedenen phylogenetischen Informationen der einzelnen Daten vereinigt werden. Allerdings bleiben die störenden Einflüsse von Insertionen, Deletionen und Rücksubstitutionen von Aminosäuren bei der phylogenetischen Rekonstruktion weiterhin erhalten.

Die in diesem Kapitel durchgeführten Untersuchungen haben neben der Erstellung der phylogenetischen Beziehungen der Spezies das Ziel, durch Kombination und Gegenüberstellung der in den vorangegangenen Kapiteln vorgestellten Methoden die Möglichkeiten zur Lokalisierung solcher störenden Sequenzdaten zu ergründen.

3.1 Untersuchungsschritte

Im ersten Schritt wurden die Datensätze mit dem Programm SORTAL sortiert (vgl. 2.1.4.2).

Der zweite Schritt beinhaltete die RASA Analyse (2.1.3). Der *tRasa* Wert wurde jeweils für den gesamten Datensatz und für jede einzelne Variabilitätsgruppe berechnet. Die Alignmentabschnitte der Variabilitätsgruppen, deren *tRasa* Wert über dem Schwellwert *tCrit* lag und somit ein signifikantes phylogenetische Signal beinhalten bildeten das RASA Alignment (vgl. 2.1.4.4).

Im folgendem dritten Schritt wurden jeweils aus den sortiertem kompletten und aus dem RASA Datensatz Bäume unter der Verwendung der Programme PROTDET und PBL (vgl. 2.1.4.3) rekonstruiert.

Bei PBL wurden mit der PROTML Option Bäume aus dem gesamten Alignment und jeweils nach Reduzierung um die höchste Variabilitätsklasse erstellt. Für das komplette Alignment erfolgte mit der *downweighting* Option des PROTDIT Programms die Berechnung von 11 Distanzmatrizen und die daraus resultierenden Baum-Spektren. Diese Option fiel bei der Analyse des RASA Alignments aufgrund der fehlenden konstanten Positionen weg.

Die besten Splits wurden jeweils aus den Spektren mit der *closest tree selection* ausgewählt.

3.2 Auswertung der Bakteriendatensätze

3.2.1 Der Archae- und Eubakteriendatensatz

Zunächst wurde der Datensatz untersucht, der sowohl Archae- und Eubakterien umfaßt (vgl. Tabelle 2.3). Durch Sortierung der Positionen mit dem Programm SORTAL ließ sich die Größe der einzelnen Variabilitätsgruppen ablesen (vgl. Tabelle 3.1). Die Ergebnisse der daraufhin durchgeführten RASA Untersuchung zeigen die Abbildungen 3.1 und 3.2.

Variabilitätsgruppe	1	2	3	4	5	6	7	8	9	10
Anzahl der Positionen	284	527	643	719	671	667	594	511	434	345
Variabilitätsgruppe	11	12	13	14	15	16	17	18	19	20
Anzahl der Positionen	223	130	133	62	45	31	34	28	32	2869

Tabelle 3.1: Übersicht der Variabilitätsgruppen des Archae-/ Eubakterienalignments

Die Abbildung 3.1 zeigt die *tRasa* Werte bezüglich der einzelnen Variabilitätsgruppen. Die Zahlen in den Klammern geben die jeweilige Anzahl der Positionen der Variabilitätsgruppen wieder. Durch die blaue Linie ist der Schwellwert *tCrit* markiert. Neben den *tRasa* Werten der Gruppen 18 und 20 liegen wider Erwarten die *tRasa* Werte bei den Gruppen mit den niedrigen Variabilitäten 2 bis 9 unterhalb des kritischen Schwellwertes *tCrit* und enthalten somit laut der RASA Definition kein eindeutiges phylogenetisches Signal. Dies steht im Gegensatz zu der Annahme, daß gerade Positionen mit niedriger Variabilität bessere phylogenetische Informationen enthalten als sehr variable Positionen.

In den Diagrammen in Abbildung 3.2 sind für jede einzelne Variabilitätsgruppe die Werte $RASA_{ij}$ über den E_{ij} Werten als Kreuze, die Ausgleichsgerade als grüne gestrichelte Linie und die Nullgerade als blaue gestrichelte Linie markiert. Bei den farblich in grün hervorgehobenen Kreuzen handelt es sich um Sequenzpaare i,j , wobei i und j für jeweils ein Archaeobakterium stehen.

Betrachtet man die Abbildungen 3.2 a bis i, die die einzelnen RASA Diagramme für die Variabilitätsgruppen 2 bis 9 zeigen, so ist zu erkennen, daß die grünen Kreuze deutlich oberhalb der Geraden und der übrigen Kreuze liegen. Dies bedeutet, daß sich selbst in den niedrigen Variabilitätsgruppen die Proteine der Archaeobakterien zum einen sehr ähnlich sind und sich zum anderen sehr stark von den übrigen Eubakterien unterscheiden.

Wie im zweiten Kapitel beschrieben, führen gerade solche langen Äste bei der RASA Analyse zu Problemen. Durch solche sehr stark abweichenden Werte wird die Steigung der Ausgleichgeraden β erhöht. Der Einfluß auf die Bestimmung der Nullgeradensteigung β_0 ist noch extremer, weil sich die Steigung aus dem Mittelwert der einzelnen Steigungen der Permutationsalignments bildet (vgl. Abschnitt 2.1.4). Dies hat zur Folge, daß β_0 einen größeren oder fast genauso großen Wert wie β annimmt und sich aus den Formeln in Abschnitt 2.1.4 für $tRasa$ ein negativer bzw. ein sehr kleiner Wert ergibt.

Die Archaeobakterien, die sich sehr stark von den Eubakterien unterscheiden, sind die bei diesem Datensatz folglich die Ursache für die niedrigen $tRasa$ Werte in den unteren Variabilitätsgruppen. Um diesem RASA Problem entgegenzutreten, blieb als einzige Möglichkeit, sich bei den folgenden Analysen nur auf die Eubakterien zu beschränken. Es wurden hierzu die Proteinalignments neu erstellt und zu einen neuen Datensatz aneinandergefügt (siehe Tabelle 2.3).

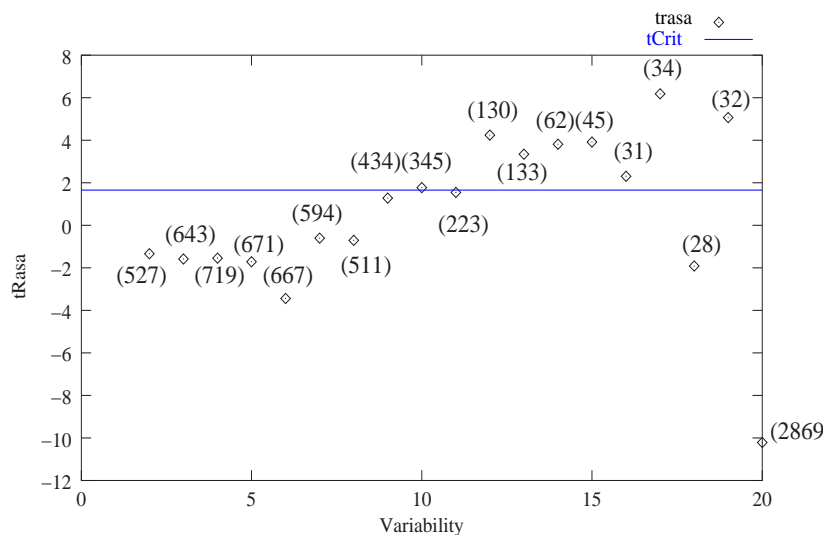


Abbildung 3.1: RASA-Analyse des Datensatzes mit Archae- und Eubakterien: Verteilung von $tRasa$ in Abhängigkeit zur Variabilität.

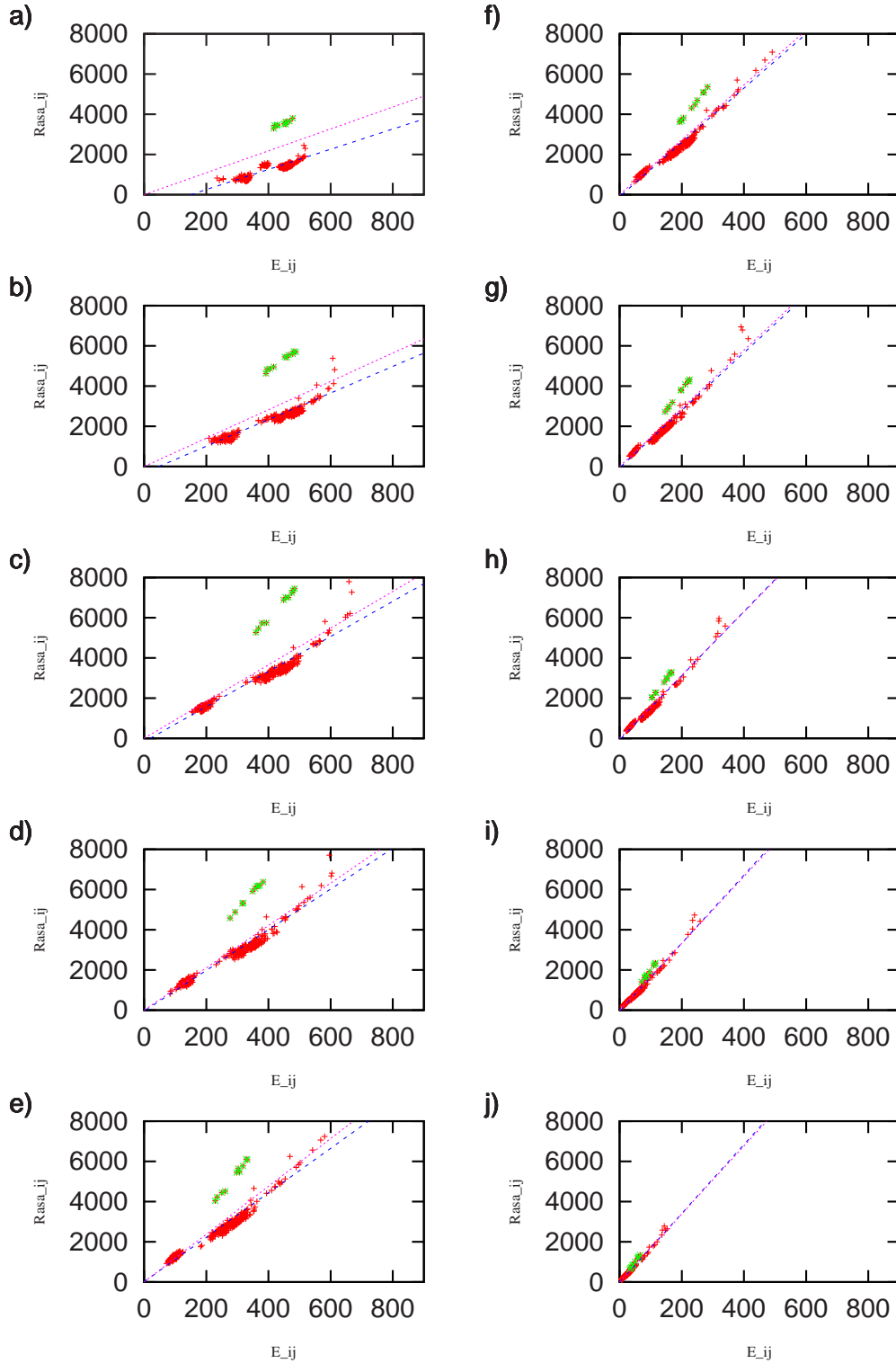


Abbildung 3.2: RASA-Analyse der Variabilitätsgruppen 2 (a) bis 11 (j) des Archae- und Eubakterien Alignments. Die Kreuze stellen die Abhängigkeit von Anzahl der Paare zu Anzahl der Triplets für alle Sequenzpaare dar. Mit den grünen Kreuze sind Sequenzpaare hervorgehoben, bei denen es sich um ein Archaeobakterienpaar handelt. Die Ausgleichsgerade ist blau, die Nullgerade pink dargestellt.

3.2.2 Der Eubakteriendatensatz

Für die in Tabelle 2.3 aufgeführten 32 Proteine wurden aus den entsprechenden Sequenzen der 21 Eubakterien neue Alignments mit CLUSTALW erstellt, aneinandergesetzt und mit dem Programm SORTAL nach Variabilitäten sortiert. Die Anzahl der Positionen betrug im Gegensatz zu dem Archae-/Eubakterien mit 8982 nur noch 8026 Positionen. Die Größen der einzelnen Variabilitätsgruppen sind Tabelle 3.2 dargestellt.

Variabilitätsgruppe	1	2	3	4	5	6	7	8	9	10
Anzahl der Positionen	679	863	830	806	678	578	493	414	291	170

Variabilitätsgruppe	11	12	13	14	15	16	17	18	19	20
Anzahl der Positionen	100	66	45	25	31	50	37	64	118	1688

Tabelle 3.2: Übersicht der Variabilitätsgruppen des Eubakterienalignments

Die Ergebnisse der RASA Analyse der einzelnen Variabilitätsklassen des Eubakteriendatensatzes geben die Abbildungen 3.3 bis 3.5 wieder. Die *tRasa* Werte im Diagramm 3.3 zeigen, daß alle Variabilitätsgruppen außer 2, 3, 13, 14, 19 und 20 den RASA Test bestehen und somit ein eindeutiges phylogenetisches Signal enthalten. Nach der Hinwegnahme des langen Archaeobakterien Astes unterlegt RASA die Annahme, daß gerade die niedrigen Variabilitätsgruppen phylogenetische Information beinhalten. Für den gesamten Datensatz weist die RASA Analyse ebenfalls das Vorhandensein eindeutiger phylogenetischer Signale auf. In den Diagrammen läßt sich auch keine eindeutige „Ausreißergruppe“ wie die der Archaeobakterien in den Abbildungen 3.2 erkennen.

Das anhand des Diagrammes 3.3 erstellte RASA Alignment enthält die Variabilitätsgruppen 4 bis 12 und 15 bis 18. Mit dem gesamten und mit dem RASA Alignment wurden wie im vorhergehenden Abschnitt beschrieben zwei Untersuchungen vorgenommen. Zum einen die PROTML Analysen von jeweils den beiden Datensätzen und nach der Reduzierung um die höchste Variabilitätsgruppe. Zum anderen erfolgte die Auswertung der beiden Datensätze mit dem Programm PROTDET. Bei dem vollständigen Datensatz wurde dabei die *downweight* Option angewendet.

Die aus den zwei Datensätzen erstellten Abstammungsbäume sind in Abbildung 3.6 zu sehen. Der Baum in 3.6a) wurde sowohl aus dem gesamten Original- als auch aus dem RASA-Alignment mit PROTML rekonstruiert. Der PROTDET Baum aus dem gesamten ursprünglichen Alignment ist in 3.6 b) und aus dem RASA Alignment in 3.6 c) zu sehen.

Nach der Reduzierung um die jeweils höchste Variabilitätsgruppe entspricht die Topologie der PROTML Untersuchung des ursprünglichen Datensatzes bis zur Reduzierung um die Variabilitätsgruppe 5 jener in 3.6 a). Bei dem RASA Alignment stimmen die ML-Topologien bis auf die letzten 4 Reduzierungs-

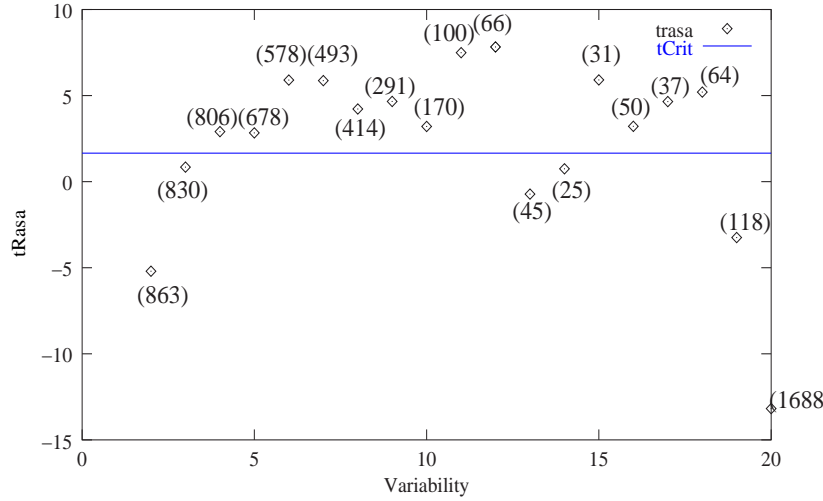


Abbildung 3.3: RASA-Analyse des Eubakterien-Datensatzes: Verteilung von $tRasa$ in Abhängigkeit zur Variabilität

schritte mit dem ML Baum in Abbildung 3.6 a) überein. Bei der *downweight* Option von PROTDET ergab sich in den 11 Schritten keine Änderung zur Topologie in Abbildung 3.6 b).

Die drei Topologien stimmen in den folgenden 6 Teilbäumen überein:

- i. **Chlamydia Gruppe:** (cp, ct)
- ii. **Spirochaeten:** (bb, tp)
- iii. (cj, hp)
- iv. (aae, tm)
- v. $((mg, mp), bs)$
- vi. $((ec, hi), vib), pae)$

Der Ast $(((((ec, hi), vib), pae), nm), xyf), rp)$ ist sowohl in den ML Bäumen sowie in dem PROTDET Baum des ursprünglichen Datensatzes enthalten. Bei dem PROTDET Baum des RASA Alignments haben lediglich *Neisseria* und *Xyella* die Positionen getauscht und *Rickettsia* ist gemeinsam mit iii.) an einem Ast zu finden. Weitere Unterschiede dieser PROTDET Topologie in Abbildung 3.6 c) zu den anderen zwei Topologien sind die Positionen den Teilbäumen i.), ii.) und iv.).

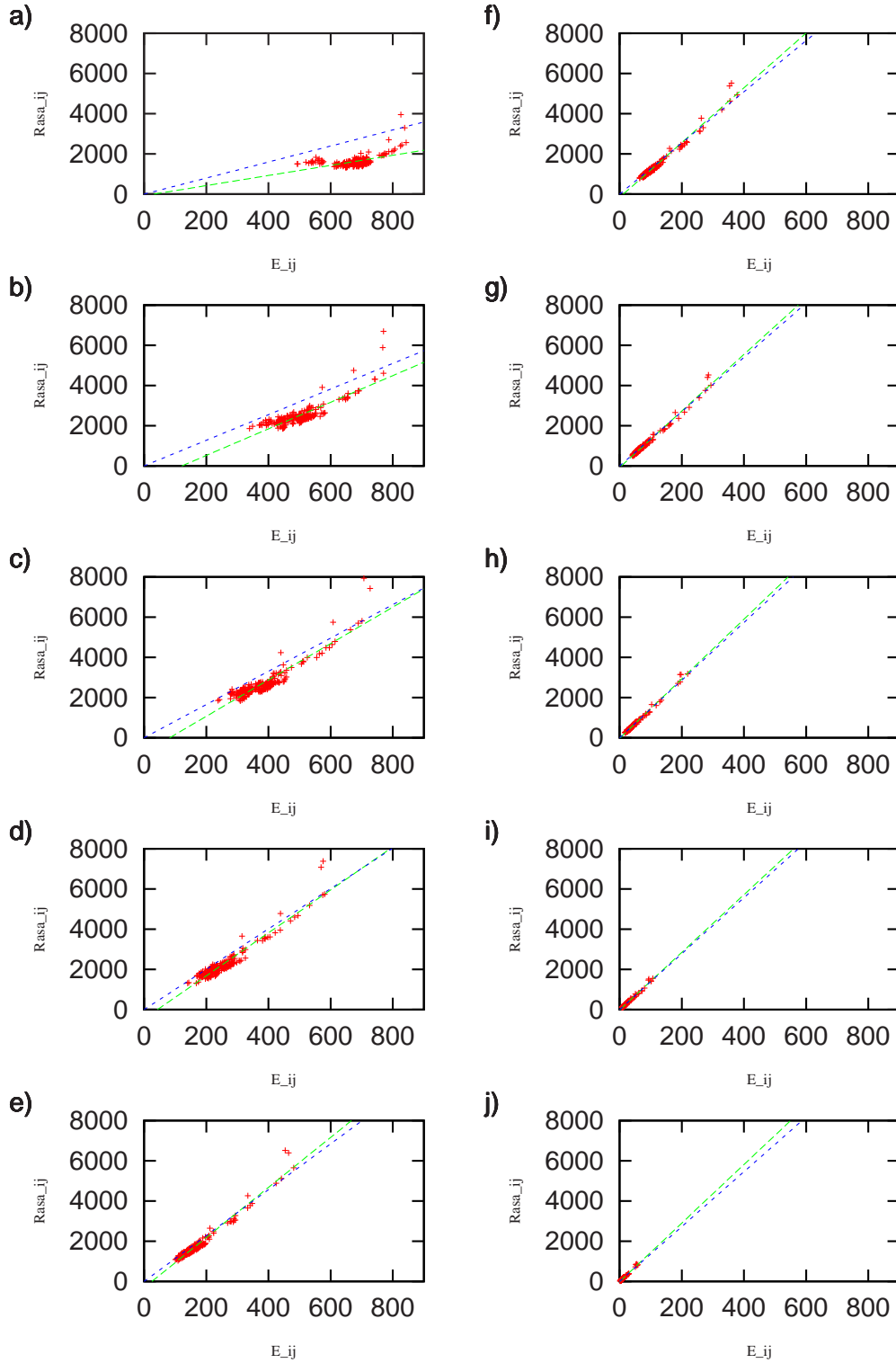


Abbildung 3.4: RASA-Analyse der Variabilitätsgruppen 2 (a) bis 11 (j) des Eubakterien Alignments. Die Kreuze stellen die Abhängigkeit von Anzahl der Paare zu Anzahl der Triplets für alle Sequenzpaare dar. Die Ausgleichsgerade ist grün, die Nullgerade blau dargestellt.

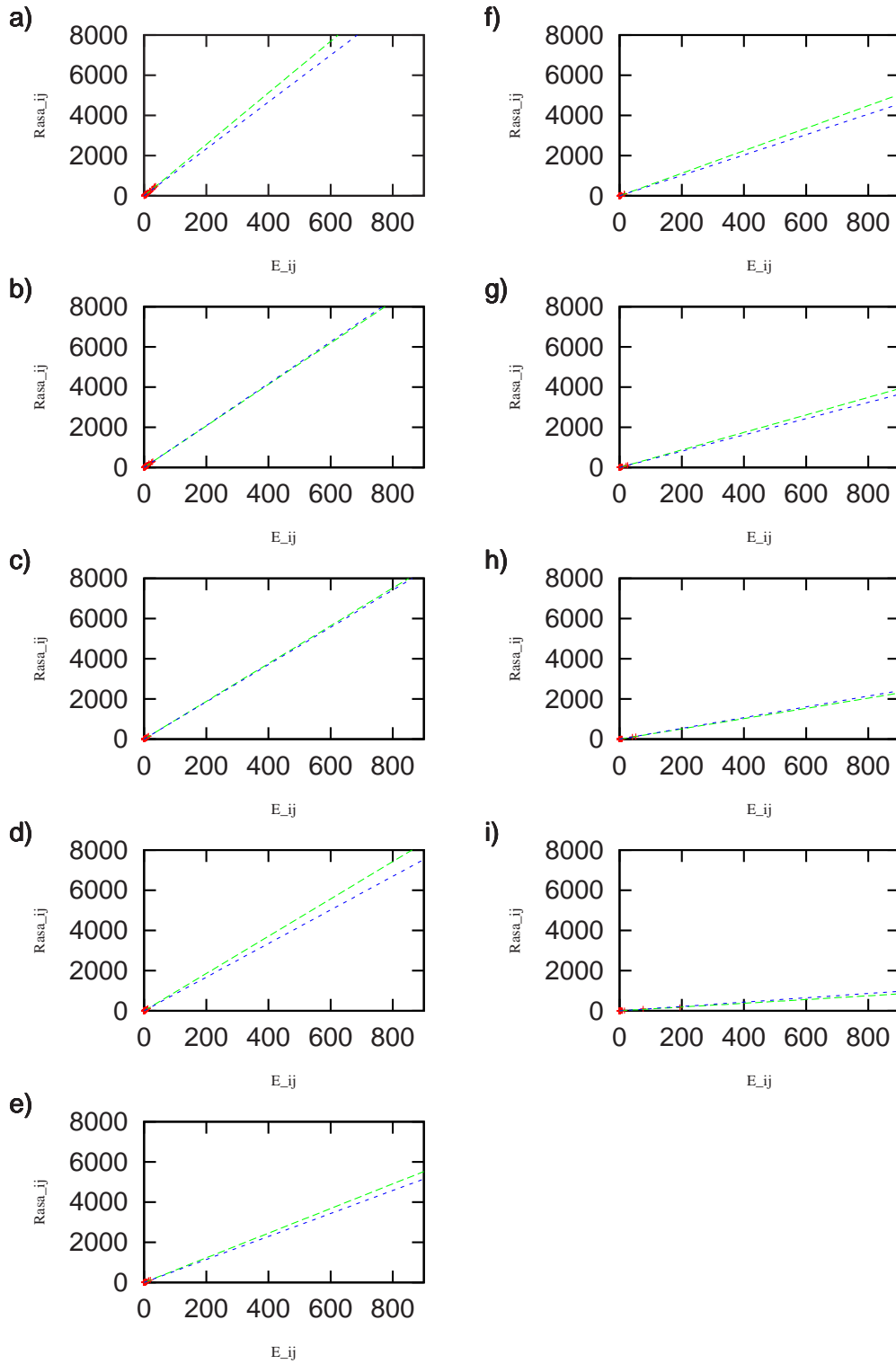


Abbildung 3.5: RASA-Analyse der Variabilitätsgruppen 12 (a) bis 20 (i) des Eubakterien Alignments. Die Kreuze stellen die Abhängigkeit von Anzahl der Paare zu Anzahl der Triplets für alle Sequenzpaare dar. Die Ausgleichsgerade ist grün, die Nullgerade blau dargestellt.

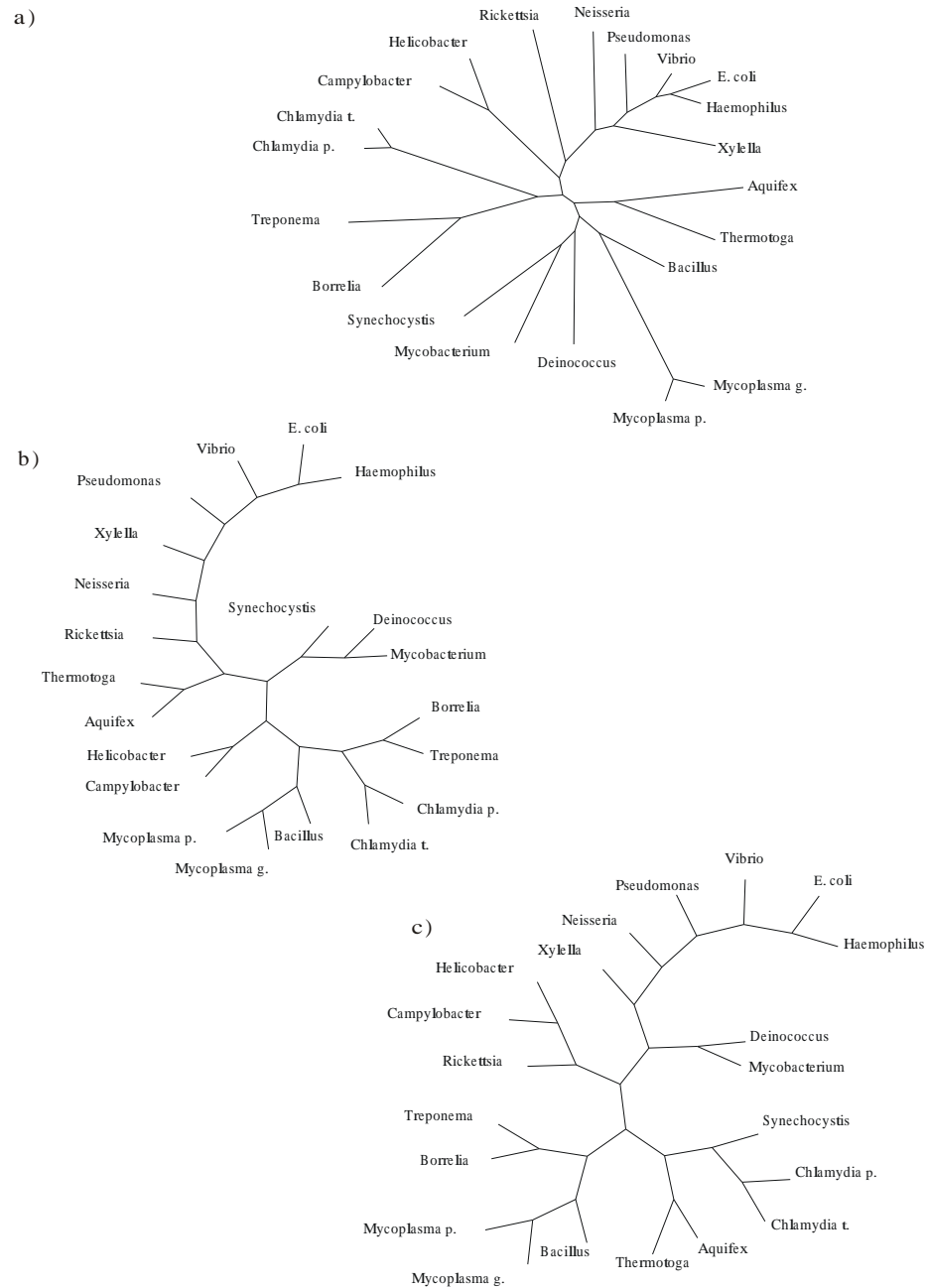


Abbildung 3.6: Phylogenetische Bäume des Eubakteriendatensatzes. a) ML Baum erstellt aus dem ursprünglichen und dem RASA Datensatz mit PROTML. b) PROTDET Baum des kompletten Datensatzes. c) PROTDET Baum des RASA Datensatzes

Sind in den Bäumen 3.6 a),b) die Chlamydia Gruppe (i) und die Spirochaeten (ii) noch am selben Ast zu finden, so bildet i.) gemeinsam mit *Synechocystis* und iv.) einen Teilbaum. Die Spirochaeten (ii) sind zusammen mit v.) an einem Teilbaum zu lokalisieren.

Neben der Plazierung des Astes iv.) unterscheiden sich die drei Topologien in den Positionen von *Deinococcus*, *Mycobacterium* und *Synechocystis*. In den ML Bäumen ist der Ast mit $((syn, mtu), der)$ enthalten. In den PROTDET Bäumen sind hingegen *Deinococcus* und *Mycobacterium* direkt benachbart. Aus der PROTDET Analyse des ursprünglichen Datensatzes ist *Synechocystis* noch am selben Ast wie dieses Paar zu finden und bei der Topologie des RASA Alignments gemeinsam mit der Chlamydia Gruppe (i), wie bereits oben erwähnt wurde.

Bei der Reduzierung um die jeweils höchste Variabilitätsgruppe erstellte PROTML bis zur Reduzierung der Variabilität 8 den oben beschriebenen Baum. In den folgenden drei Schritten beinhalten die jeweiligen Topologien anstatt $((syn, mtu), der)$ den Ast $((der, mtu), mtu)$. In den letzten drei Schritten wechseln mit Ausnahme der Teilbäume $((((ec, hi), vib), pae), xyf)$, i), ii), iii), iv) sowie der Mycoplasma Gruppe alle Äste ihre Positionen.

Bei der ML Analyse des RASA Alignments ergab bis zum 9. Reduzierungsschritt, die Reduzierung um Variabilitätsgruppe 8, die Topologie in 3.6 a). In den folgenden drei Schritten änderten nur *Deinococcus*, *Mycobacterium* und *Synechocystis* ihre Positionen. Nach dem letzten Schritt beinhaltete das Alignment nur noch die Variabilitätsgruppe 3. Der daraus erstellte ML Baum enthielt von der beschriebenen ML Topologie nur noch die Äste i) - v) und $((((ec, hi), vib), pae), xyf)$.

3.2.2.1 Qualitativer Vergleich der erstellten Eubakterienbäume

Wie im vorhergehenden Abschnitt beschrieben, wurden mit den Programmen PROTDET und PROTML für den gesamten und den RASA Datensatz drei verschiedene Topologien erstellt (vgl. Abbildung 3.6). Die Topologien der der ML Bäume waren dabei für beide Datensätze identisch (vgl. Abbildung 3.6 a)). Die Topologien der PROTDET Bäume des gesamten und des RASA Datensatzes unterschieden sich untereinander und wiesen zudem Differenzen zur ML Topologie auf (vgl. Abbildung 3.6 b) und c)).

Die drei Topologien werfen nun zum einen die Frage auf, ob sich die Reduzierung des Datensatzes mit Hilfe der Variabilitätseinteilung und der RASA Methode auf das RASA Alignment wirklich bei der Lokalisierung phylogenetisch informativer Daten eignet. Zum anderen ist auch die Qualität der Topologien in Abhängigkeit der Rekonstruktionsmethoden zu berücksichtigen.

Zunächst soll die ML Topologie betrachtet werden. Ausgehend von einer NJ Topologie wurde mit dem *Local Rearrangement* Prinzip der ML Baum

erstellt. Bei diesem Vorgehen wird für die entsprechenden Teilbäume eine lokale Wahrscheinlichkeit berechnet. Diese wird auch LBP (Local Bootstrap Probability) genannt. Der entgültige ML Baum setzt sich aus den wahrscheinlichsten Teilbäumen zusammen.

Die zwei ML Topologien sind zwar identisch, allerdings unterscheiden sich die LBPs für einige Teilbäume. Die Tabelle 3.3 gibt die LBPs der einzelnen Teilbäume beider ML Bäume wieder.

Alle Äste, die in der Topologie des gesamten Datensatzes mit den LBP von 100% bzw. knapp 100% rekonstruiert werden, haben auch bei der Topologie des RASA Alignments einen Wert von 100% bzw. knapp 100%. Mit einer Ausnahme gehören die Äste zu den „sicheren“ Ästen, an denen Teilbäume mit einem OTU Paar liegen.

Die Ausnahme (mtu, syn) stellt mit 58% allerdings deutlich einen „kritischen“ Ast dar. Bei der Rekonstruktion des ML Baums aus dem RASA Datensatz weißt dieser Ast einen „sicheren“ Wert von 76 % auf. Dafür sinkt aber der Ast $((mtu, syn), der)$ von 94% deutlich auf nur 75%. Ebenfalls verschlechterte sich der Ast $((((mg, mp), bs), ((mtu, syn), der)), (aae, tm))$ eindeutig von 81% auf 69%.

Die Positionierung von *Synechocystis*, *Deinococcus* und *Mycobacterium* stellt wie bereits beschrieben einen markanten Unterschied zwischen den ML und den PROTDET Topologien dar. So bildet in den PROTDET Bäumen *Deinococcus* und *Mycobacterium* ein direktes OTU Paar. Die NJ Topologie, die als Ausgangstopologie für das *Local Rearrangement* mit PROTML verwendet wurde, weißt ebenfalls dieses OTU Paar auf. Allerdings ergeben sich für dieses Paar sowohl für das ursprüngliche sowie für das RASA Alignment deutlich geringere LBP Werte (bei den jeweils kompletten Datensätzen jeweils nur 15 %).

Diese beiden möglichen Äste lassen sich bei PROTDET anhand der berechneten Frequenzen vergleichen. In Tabelle 3.4 sind alle Splits, die in dem beiden PROTDET und in den PROTML Analysen erstellt worden sind, aufgeführt. Die Frequenzen der Splits sind in den Diagrammen 3.7 a) bezüglich des gesamten und 3.7 b) bezüglich des RASA Datensatzes dargestellt.

Im Vergleich der Diagramme ist gut zu erkennen, daß die Frequenzen der Splits des PROTDET Baumes aus dem RASA Alignment deutlich höher sind, als die des gesamten Datensatzes. Auch die Splits 13, 14, 16 weisen zum Teil deutlich größere Frequenzen bei der PROTDET Analyse des RASA Datensatzes auf als bei der des gesamten Datensatzes, obwohl sie in der Topologie nicht vorhanden sind (vgl. Abbildung 3.7 und Tabelle 3.4). Dafür ergeben sich für die Splits 26, 27 und 28, die nur in der ML Topologie zu finden sind niedrigere Frequenzen.

Um die Positionierung von *Synechocystis*, *Deinococcus* und *Mycobacterium* anhand der PROTDET Frequenzen vergleichen zu können, müssen die Splits

	LBP - gesamte Daten	LBP - RASA Daten
(aae, tm)	100	100
(mg, mp)	100	100
$((mg, mp), bs)$	100	99
(mtu, syn)	58	76
$((mtu, syn), der)$	94	75
$((((mg, mp), bs), ((mtu, syn), der)))$	96	98
$(((((mg, mp), bs), ((mtu, syn), der)), (aae, tm)))$	81	69
(bb, tp)	100	100
(cp, ct)	100	100
$((bb, tp)(cp, ct))$	100	98
$(((((mg, mp), bs), ((mtu, syn), der)), (aae, tm)), ((bb, tp)(cp, ct)))$	98	98
(cj, hp)	100	100
(ec, hi)	100	100
$((ec, hi), vib)$	100	100
$((((ec, hi), vib), pae))$	100	100
$(((((ec, hi), vib), pae), xyf),)nm)$	100	96
$((((((ec, hi), vib), pae), xyf),)nm), rp)$	100	100

Tabelle 3.3: Die ML Topologie umfaßt 17 Äste. Die LBP Werte dieser Äste sind für den gesamten Datensatz in der mittleren und die des RASA Datensatzes in der dritten Spalte aufgeführt.

15, 26, 12 und 19 betrachtet werden. Für den von den der ML Analyse erstellten Split 26 (mtu, syn) ergeben sich gegenüber dem PROTDET Split (der, mtu) wesentlich niedrigere Frequenzen. Im Falle des RASA Alignments ist die Frequenz des Splits 26 sogar negativ (3.7 b)). Die Positionierung von *Synechocystis* am *Chlamydia* Ast (Splits 19) erhält nur bei der PROTDET Topologie des RASA Datensatzes eine deutlich größere Frequenz als Split 12.

#	Split
1	(<i>mg, mp</i>)
2	(<i>cp, ct</i>)
3	(<i>cj, hp</i>)
4	(<i>bb, tp</i>)
5	(<i>ec, hi, nm, pae, vib, xyf</i>)
6	(<i>ec, hi, vib</i>)
7	(<i>aae, tm</i>)
8	(<i>ec, hi, pae, vib</i>)
9	(<i>ec, hi, nm, pae, rp, vib, xyf</i>)
10	(<i>bs, mg, mp</i>)
11	(<i>ec, hi</i>)
12	(<i>der, mtu, syn</i>)
13	(<i>bb, cp, ct, tp</i>)
14	(<i>ec, hi, pae, vib, xyf</i>)
15	(<i>der, mtu</i>)
16	(<i>bb, bs, cp, ct, mg, mp, tp</i>)
17	(<i>bb, bs, cj, cp, ct, hp, mg, mp, tp</i>)
18	(<i>aae, ec, hi, nm, pae, rp, tm, vib, xyf</i>)
19	(<i>cp, ct, syn</i>)
20	(<i>ec, hi, nm, pae, vib</i>)
21	(<i>bb, bs, mg, mp, tp</i>)
22	(<i>cj, hp, rp</i>)
23	(<i>aae, bb, bs, cp, ct, mg, mp, syn, tm, tp</i>)
24	(<i>(aae, cp, ct, syn, tm)</i>)
25	(<i>der, ec, hi, mtu, nm, pae, vib, xyf</i>)
26	(<i>mtu, syn</i>)
27	(<i>bs, der, mg, mp, mtu, syn</i>)
28	(<i>aae, bs, der, mg, mp, mtu, syn, tm</i>)
29	(<i>aae, bb, bs, der, mg, mp, mtu, syn, tm, tp</i>)

Tabelle 3.4: Die Splits 1 bis 18 sind im PROTDET Baum des gesamten Eubakteriendatensatzes enthalten. Der PROTDET Baum des RASA Alignments setzt sich aus den Splits 1 bis 8, 10, 11, 15, 19 bis 25 zusammen. Die Splits 26 bis 29 sind in der ML Topologie, aber nicht in den PROTDET Topologien enthalten. Die Zuordnung der Splits zu den Nummern erfolgte bezüglich der Diagramme in 3.7

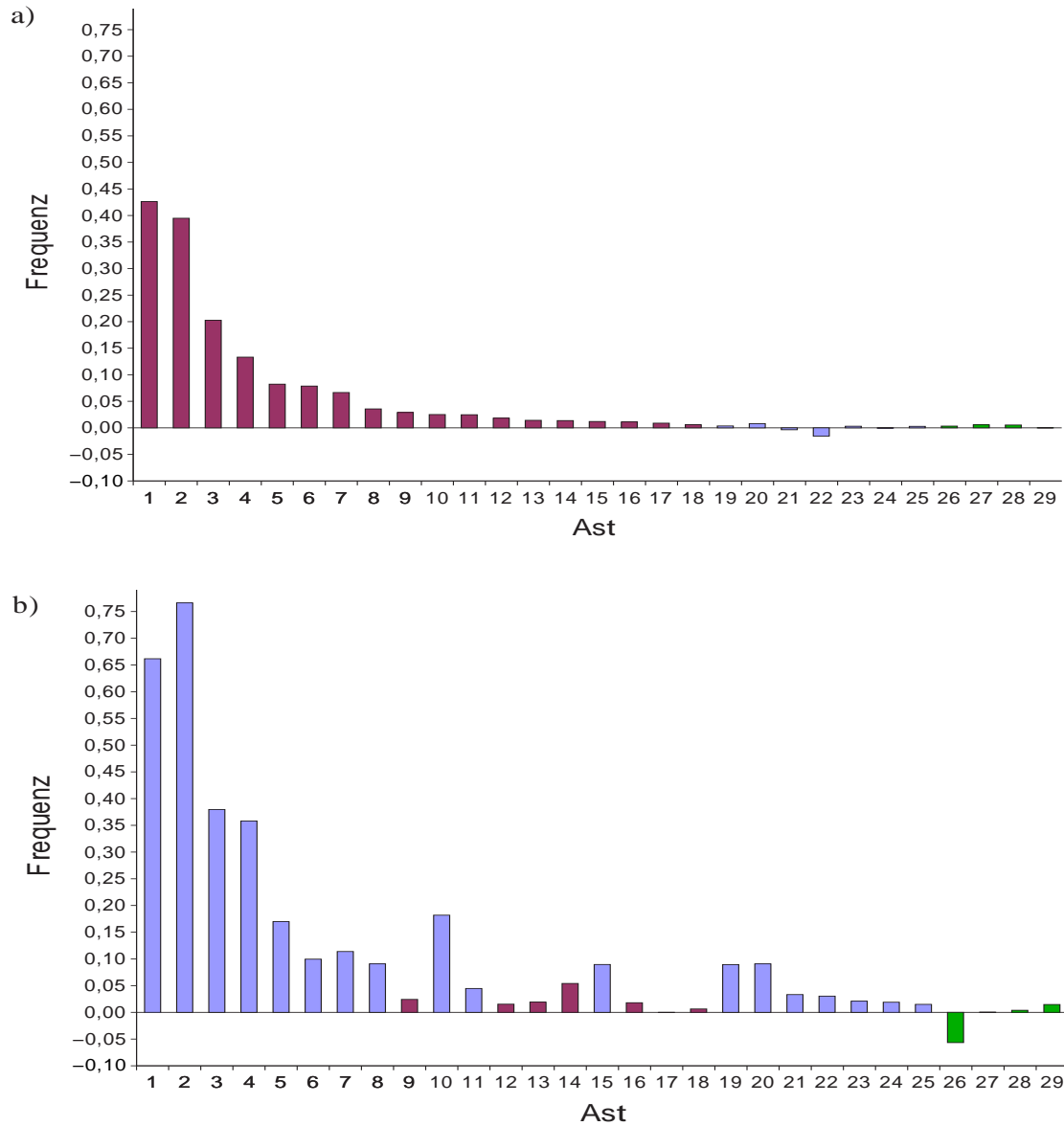


Abbildung 3.7: Die Frequenzen der Splits, die in Tabelle 3.4 aufgeführt sind, wurden mit PROTDET für das gesamte (a) und das RASA Alignment (b) erstellt. Die Splits des gesamten Datensatzes sind in (a) und (b) rot und die des RASA Datensatzes lila markiert. Die grünen Balken der Splits 26 bis 29 stehen für die Frequenzen der Splits, die in der ML Topologie nicht aber in den PROTDET Topologien enthalten sind.

Durch Berechnung des Likelihoods für jede der drei Topologien lassen sich die Bäume durch statistische Tests wie dem Kishino-Hasegawa Test (Kishino and Hasegawa (1989)) vergleichen. Mit dem Programm PROTML kann der Test von Kishino-Hasegawa durchgeführt werden.

Der Vergleich der drei Topologien führte zu folgenden Ergebnissen.

protml 2.3b3 JTT 3 trees 21 OTUs 8026 sites.

Tree	ln L	Diff ln L	S.E.	#Para	AIC	Diff AIC	TBL	RELL-BP
1	-195877.3	0.0	<-best	39	391832.5	0.0	ME	1.0000
2	-196209.5	-332.3	53.0	39	392497.1	664.6	5.5	0.0000
3	-196449.6	-572.4	84.9	39	392977.2	1144.7	7.6	0.0000

protml 2.3b3 JTT 3 trees 21 OTUs 3778 sites.

Tree	ln L	Diff ln L	S.E.	#Para	AIC	Diff AIC	TBL	RELL-BP
1	-142670.8	0.0	<-best	39	285419.7	0.0	ME	1.0000
2	-142865.8	-194.9	35.2	39	285809.6	389.9	11.3	0.0000
3	-143028.0	-357.2	55.2	39	286134.1	714.4	17.7	0.0000

Bei den Bäumen handelt es sich bei Baum 1 um die ML Topologie, bei 2 um die PROTDET Topologie aus dem gesamten und bei 3 um die PROTDET Topologie des RASA Datensatzes. Entscheidend bei dem Test ist der Wert $\ln L$, der natürliche Logarithmus aus dem berechneten Likelihood Wert. Dieser Wert ist sowohl beim Test des gesamten wie auch beim RASA Alignment am geringsten. Folglich ergibt sich für Baum 1 der größte Likelihood. Auch für den Wert AIC (Akaike Information Criterium), mit $AIC = -2 \ln L + 2 \#Para$, ergibt Baum 1 das beste Resultat. Hierbei gibt der Parameter $\#Para$ die Anzahl aller internen und externen Äste des Baumes an.

3.3 Auswertung des Plastidendatensatzes

Mit den in den vorangegangenen Abschnitten beschriebenen Methoden wurde nun auch der Plastidendatensatz (2.6) untersucht. Begonnen wurde wiederum mit der Sortierung nach Variabilitätsgruppen mit dem Programm SORTAL (vgl. Tabelle 3.5) gefolgt von der RASA Analyse, deren Ergebnisse in den Diagrammen der Abbildungen 3.8 bis 3.10 festgehalten sind.

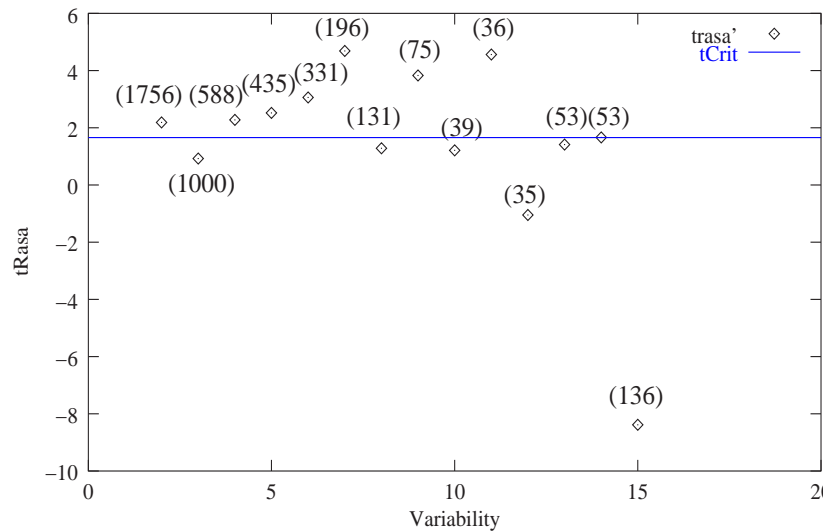
In der Darstellung der $tRasa$ Werte in Diagramm 3.8 ist zu erkennen, daß $tRasa$ der Variabilitätsgruppen 3, 8, 10, 12, 13 und 15 unterhalb des Schwellwertes $tCrit$ liegen, also nach dem RASA Test kein eindeutiges phylogenetisches Signal beinhalten. Für den gesamten Datensatz liegt der $tRasa$ knapp unterhalb des Schwellwertes.

Variabilitätsgruppe	1	2	3	4	5	6	7	8	9	10
Anzahl der Positionen	3155	1756	1000	588	435	331	196	131	75	39

Variabilitätsgruppe	11	12	13	14	15	16
Anzahl der Positionen	36	35	53	53	136	289

Tabelle 3.5: Übersicht der Variabilitätsgruppen des Plastidenalignments

Das aus dem Plastidendatensatz erstellte RASA Alignment beinhaltet somit die Variabilitätsgruppen 2, 4, 5, 6, 7, 9, 11 und 14 und umfaßt 3470 Positionen.

Abbildung 3.8: RASA-Analyse des Plastidendatensatz: Verteilung von $tRasa$ in Abhängigkeit zur Variabilität

Der ursprüngliche und der RASA Plastidendatensatz wurden ebenso wie der Eubakteriendatensatz mit den Programmen PROTDET und PROTML untersucht. Die PROTML Analysen wurden mit PBL nach der Reduzierung der jeweils höchsten Variabilitätsgruppe ausgeführt. Die *downweight* Option von PROTDET konnte wiederum nur für den ursprünglichen Datensatz angewandt werden, weil dieser im Gegensatz zum RASA Alignment konstante Positionen enthält.

Die ML Topologien, die sowohl für das ursprüngliche und für das RASA Alignment erstellt wurde, sind in Abbildung 3.11a) und b) zu sehen. Die Topologie unterscheiden sich lediglich durch die Position von *Cyanophora*. Im ML Baum des ursprünglichen Datensatzes bildet es mit *Synechocystis* einen

Ast (3.11a)). Bei der Topologie des RASA Alignments bilden *Cyanophora* und *Synechocystis* keinen gemeinsamen Ast mehr, sind aber trotzdem direkt benachbart. Bei der Reduzierung der jeweils höchsten Variabilitätsgruppe bleibt diese Topologie für den ursprünglichen Datensatz bis zu den letzten 3 und beim RASA Alignment bis zu den letzten 5 Reduzierungsschritten erhalten. Die PROTDET Bäume sind in den Abbildungen 3.11 c), für den ursprünglichen und 3.11d) für das RASA Alignment dargestellt. In den 11 *downweight* Schritten ändert sich an der Topologie 3.11 c) nichts.

Die vier Topologien stimmen in drei Ästen überein.

- i. (((((mai, ory), (nic, oen)), pin), mar), mes)
- ii. (cya, syn)
- iii. (gui, por)

In allen vier Bäumen sind die Grünalgen *Nephroselmis* und *Chlorella* sowie *Euglena* direkt zu Ast i. benachbart. Allerdings unterscheiden sich die Topologien in der Position von *Euglena*. In den ML Topologien ist *Euglena* am Ast der direkt benachbarten Grünalgen zu finden. Im PROTDET Baum des ursprünglichen Datensatzes bilden *Euglena* und die Grünalgen getrennte Äste. Im Fall des RASA Alignments bildet *Euglena* zwar wieder gemeinsam mit den Grünalgen einen Ast, allerdings direkt benachbart zu *Chlorella*.

In allen vier Topologien sind *Cyanidium* und *Odontella* jeweils benachbart, wobei sie aber in den ML Bäumen gemeinsam einen Ast bilden , bei den PROTDET Topologien jeweils einzelne Äste bilden. Zudem unterscheiden sich Positionen in den zwei PROTDET Bäumen von denen des ML Baumes.

Bei der Reduzierung der jeweils höchsten Variabilitätsstufe des ursprünglichen Datensatzes ergeben sich bei den Topologien zu der in Abbildung 3.11a) bis zur Reduzierung um die Variabilitätsgruppe 7 keine Veränderungen. In den folgenden 3 Schritten ändert sich lediglich die Positionen von *Euglena* und *Pinus*. Nach dem letzten Reduzierungsschritten tauschen zudem noch *Odontella* und *Guillardia* ihre Positionen und *Mesostigma* bildet mit *Nephroselmis* und *Chlorella* mit *Euglena* je einen Ast.

Beim RASA Datensatz ergibt sich bis zum 7. Reduzierungsschritt keine Änderung gegenüber dem ML Baum (3.11b)). Im folgenden Schritt bildet *Mesostigma* mit *Euglena* und den Grünalgen gemeinsam den Ast (((chl, nep), eug), mes). Nach dem nächsten Schritt ist *Mesostigma* wieder an der Position wie in 3.11b) zu finden. Allerdings tauschen *Euglena* und *Nephroselmis* ihre Positionen. Außerdem stellen in nur dieser Topologie *Nicotiana* und *Oenothera* keinen gemeinsamen Ast sondern zwei benachbarte einzelne Äste dar. Im nächsten Schritt bildet zunächst *Mesostigma* mit *Nephroselmis* einen Ast. In letzten 2 Topologien ergänzt *Nephroselmis* Ast i. In diesen eben beschriebenen 5 Schritten haben ebenfalls *Odontella* und *Guillardia* ihre Positionen bezüglich 3.11b) gewechselt.

3.3.1 Qualitativer Vergleich der Topologien

Die vier erstellten Topologien des Plastidendatensatzes weisen weniger Unterschiede zueinander auf, als die Bakterienstammbäume (3.6). Um den Einfluß der Einteilung in Variabilitätsklassen sowie der RASA Analyse auf die Qualität der Topologien besser zu veranschaulichen, sollen wiederum die LBP bzgl. der PROTML (Tabelle 3.6) und die Frequenzen bzgl. der PROTDET Analysen (Abbildung 3.12) herangezogen werden.

Wie bereits beschrieben unterscheiden sich die zwei ML Topologien lediglich in der Positionierung von *Cyanophora* und *Synechocystis*. Die LBP Werte dieser zwei Äste in den letzten zwei Zeilen der Tabelle 3.6 weisen mit 56% bzw. 54% fast gleich geringe und somit wenig aussagekräftige Werte auf. Auch in den übrigen LBP gibt es zwischen den gesamten und dem RASA Daten kaum Unterschiede. Lediglich die benachbarte Position von *Mesostigma* zu den Landpflanzen liegt mit 61% deutlich unter der LBP von 92 % des gesamten Datensatzes. Auffallend ist der LBP Wert des Astes (*chl, nep*), der gegenüber den übrigen Ästen für beide Datensätze deutlich geringere LBP Werte aufweist.

	LBP - gesamte Daten	LBP - RASA Daten
(<i>can, odo</i>)	99	99
(<i>gui, por</i>)	100	98
((<i>can, odo</i>), (<i>gui, por</i>))	100	100
(<i>chl, nep</i>)	54	61
((<i>chl, nep</i>), <i>eug</i>)	75	75
(<i>mai, ory</i>)	100	100
(<i>nic, oen</i>)	100	100
((<i>mai, ory</i>), (<i>nic, oen</i>))	100	100
(((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>)	100	100
(((((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>), <i>mar</i>)	100	100
((((((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>), <i>mar</i>), <i>mes</i>)	92	61
(((<i>chl, nep</i>), <i>eug</i>), ((((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>), <i>mar</i>), <i>mes</i>))	100	100
(((<i>can, odo</i>), (<i>gui, por</i>)), ((<i>chl, nep</i>), <i>eug</i>), ((((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>), <i>mar</i>), <i>mes</i>)))	56	-
(<i>cya</i> , (((<i>chl, nep</i>), <i>eug</i>), ((((<i>mai, ory</i>), (<i>nic, oen</i>)), <i>pin</i>), <i>mar</i>), <i>mes</i>)))	-	54

Tabelle 3.6: Beide ML Topologien umfassen jeweils 13 Äste. Die unteren beiden Äste sind jeweils nur in der ML Topologie des gesamten oder des RASA Datensatzes vorhanden. Die LBP Werte dieser Äste sind für den gesamten Datensatz in der mittleren und die des RASA Datensatzes in der dritten Spalte aufgeführt.

In den Frequenzen der PROTDET Topologien in den Diagrammen 3.12 bezüglich der Splits aus Tabelle 3.7 ist wie auch bei den Analysen der Bakterienriendaten festzustellen, daß die Frequenzen des RASA Datensatzes wesentlich

größer sind als die der gesamten Daten. Sogar die Splits 10 und 13, die lediglich im gesamten, nicht aber im RASA Datensatz enthalten sind, weisen bei der PROTDDET Analyse des RASA Alignments leicht höhere Frequenzen gegenüber denen des Gesamtalignments auf.

Für den Split 16 (*can, odo*) der in beiden ML Topologien enthalten ist, ergeben beide PROTDDET Analyse negative Frequenzen. Laut der PROTDDET Interpretation bedeuteten diese Werte, daß die Daten diesem Split widersprechende Informationen enthalten. Allerdings sind durch die Splits 7 und 12 *Cyanidium* und *Odontella* in den PROTDDET Topologien ebenfalls direkt benachbart. Lediglich die Positionierung an einem gemeinsamen Ast geht aus den PROTDDET Untersuchungen nicht hervor. Für den Split 17 des ML Baums des RASA Alignments ergibt sich bei den PROTDDET Analyse des RASA Datensatzes eine deutlich höhere Frequenz als für den Gesamtdatensatz. Die PROTDDET Untersuchung weist somit ebenfalls ein deutliches Signal im RASA Datensatz nach, daß *Synechocystis* und *Cyanophora* nicht gemeinsam an einem Ast plziert sind.

#	Split
1	(<i>mai, mar, nic, oen, ory, pin</i>)
2	(<i>mai, ory</i>)
3	(<i>mai, nic, oen, ory</i>)
4	(<i>can, cya, gui, odo, por, syn</i>)
5	(<i>mai, nic, oen, ory, pin</i>)
6	(<i>cya, syn</i>)
7	(<i>can, gui, odo, por</i>)
8	(<i>gui, por</i>)
9	(<i>nic, oen</i>)
10	(<i>can, cya, eug, gui, odo, por, syn</i>)
11	(<i>mai, mar, mes, nic, oen, ory, pin</i>)
12	(<i>gui, odo, por</i>)
13	(<i>chl, nep</i>)
14	(<i>chl, eug</i>)
15	(<i>chl, eug, nep</i>)
16	(<i>can, odo</i>)
17	(<i>can, gui, odo, por, syn</i>)

Tabelle 3.7: Die Splits 1 bis 13 sind im PROTDDET Baum des gesamten Plastidendatensatzes enthalten. Der PROTDDET Baum des RASA Alignments setzt sich aus den Splits 1 bis 9, 11, 14 und 15 zusammen. Der Split 16 ist in beiden ML Topologien, der Split 17 in der Topologie des RASA Datensatzes, aber nicht in den PROTDDET Topologien enthalten. Die Zuordnung der Splits zu den Nummern erfolgte bezüglich der Diagramme in 3.12

Um die 4 Topologien vergleichen zu können, wurde wiederum der Kishino-Hasegawa-Test (Kishino and Hasegawa (1989)) mit PROTML für den gesamten und für den RASA Datensatz durchgeführt.

Es ergaben sich für die Topologien 1 , ML Baum aus gesamten Datensatz, 2, ML Baum aus dem RASA Datensatz, 3, PROTDET Baum aus dem gesamten Datensatz, und 4, PROTDET Baum aus dem RASA Datensatz, die folgenden Resultate.

protml 2.3b3 JTT 4 trees 16 OTUs 8308 sites.

Tree	ln L	Diff ln L	S.E.	#Para	AIC	Diff AIC	TBL	RELL-BP
1	-116220.4	0.0	<-best	29	232498.7	0.0	ME	0.5984
2	-116229.4	-9.0	25.3	29	232516.7	18.0	0.2	0.3596
3	-116295.9	-75.5	37.5	29	232649.8	151.0	0.4	0.0155
4	-116284.5	-64.1	35.7	29	232627.0	128.3	0.6	0.0265

protml 2.3b3 JTT 4 trees 16 OTUs 3470 sites.

Tree	ln L	Diff ln L	S.E.	#Para	AIC	Diff AIC	TBL	RELL-BP
1	-69517.9	-2.1	14.5	29	139093.9	4.2	ME	0.4225
2	-69515.9	0.0	<-best	29	139089.7	0.0	0.2	0.5536
3	-69559.2	-43.4	26.1	29	139176.5	86.7	1.4	0.0160
4	-69562.6	-46.7	25.4	29	139183.2	93.4	1.9	0.0079

Im Fall des gesamten Alignment stellte sich auch die ML Topologie des gesamten Datensatzes als die beste heraus. Für den RASA Datensatz ergab sich als die beste Topologie die ML Topologie des RASA Alignments.

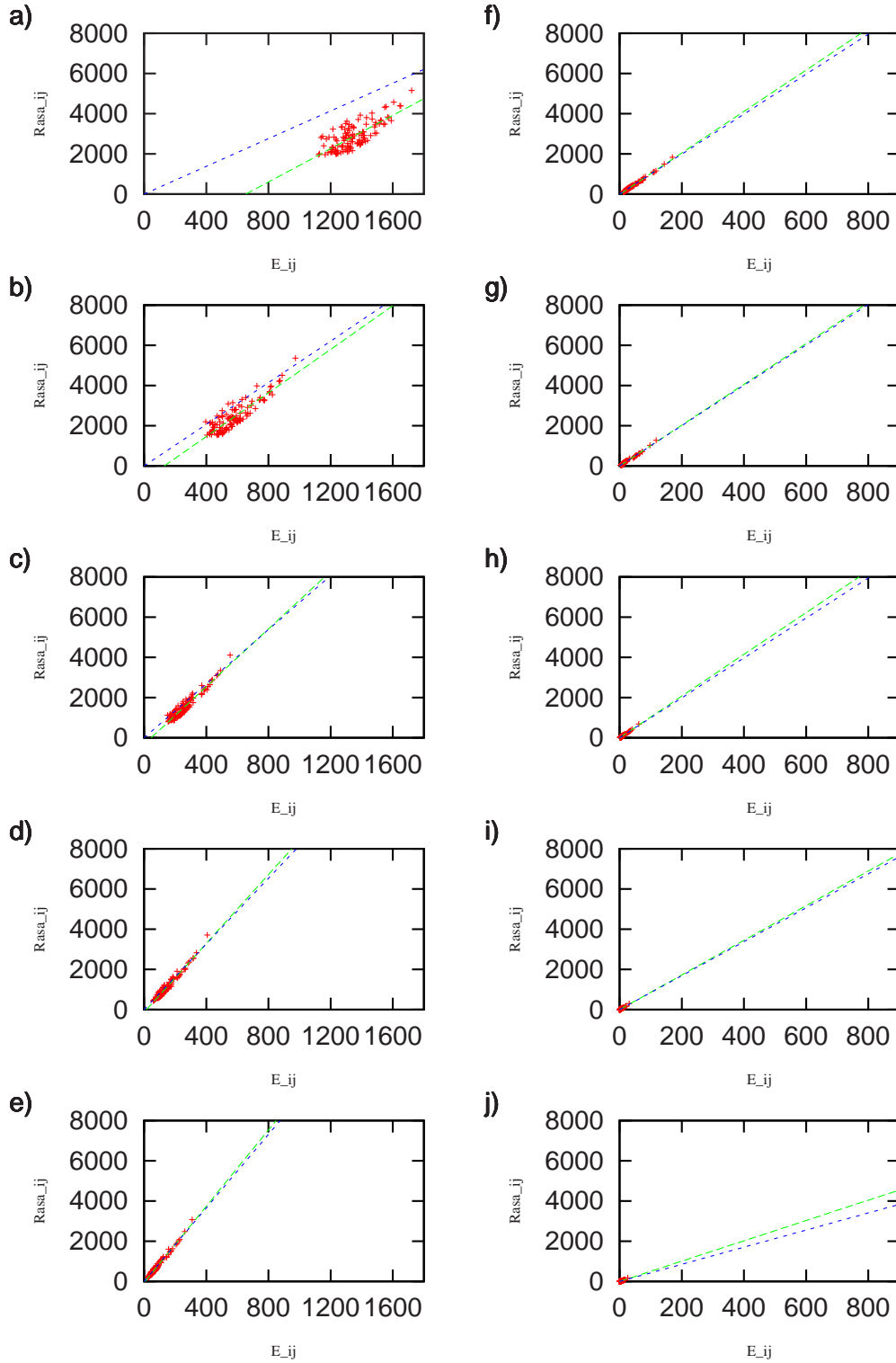


Abbildung 3.9: RASA-Analyse der Variabilitätsgruppen 2 (a) bis 11 (j) des Plastiden Alignments. Die Kreuze stellen die Abhängigkeit von Anzahl der Paare zu Anzahl der Triplets für alle Sequenzpaare dar. Die Ausgleichsgerade ist grün, die Nullgerade blau dargestellt.

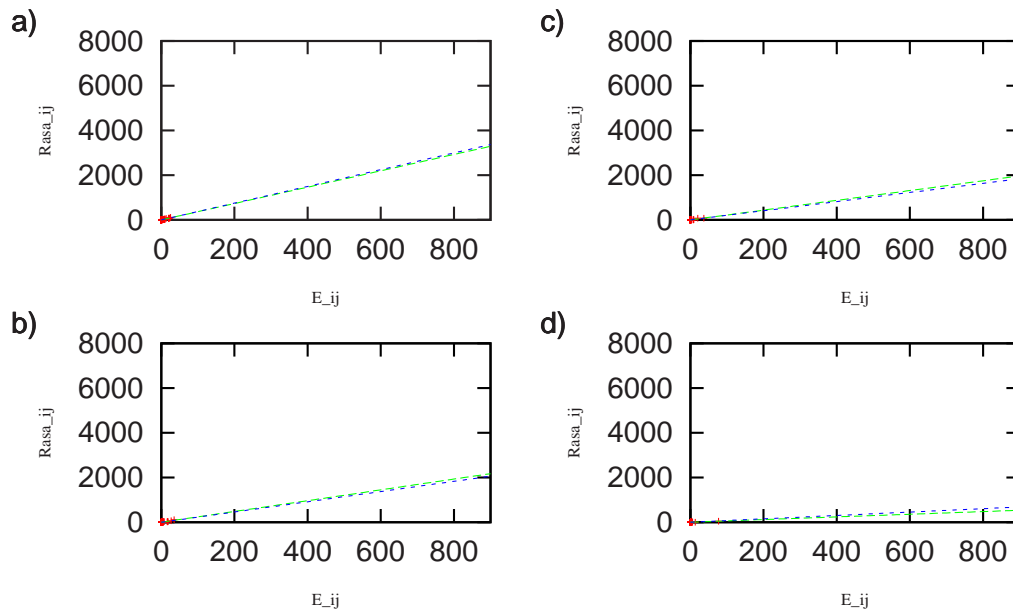


Abbildung 3.10: RASA-Analyse der Variabilitätsgruppen 12 (a) bis 14 (d) des Plastiden Alignments. Die Kreuze stellen die Abhängigkeit von Anzahl der Paa-re zu Anzahl der Triplets für alle Sequenzpaare dar. Die Ausgleichsgerade ist grün, die Nullgerade blau dargestellt.

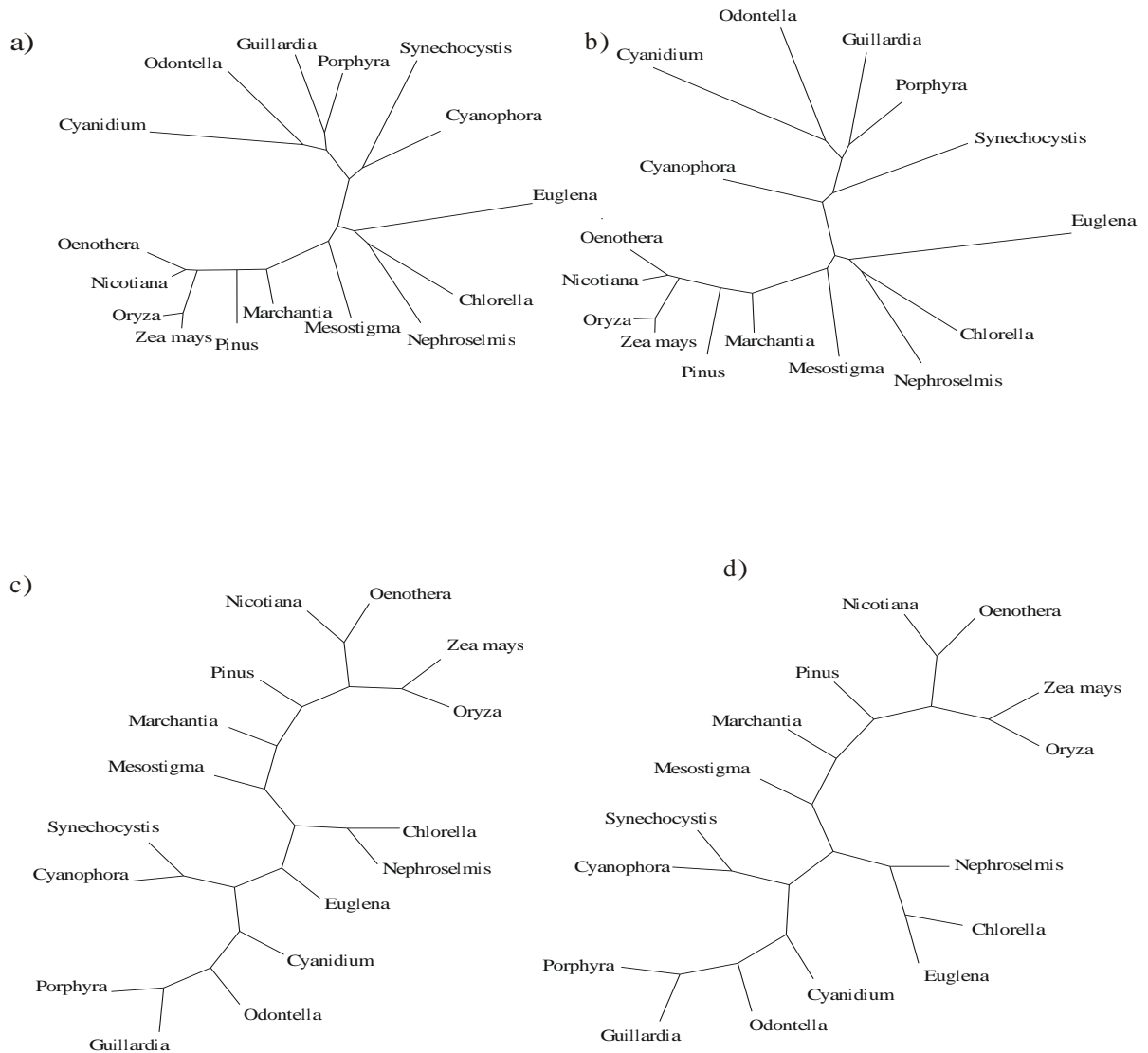


Abbildung 3.11: Phylogenetische Bäume des Plastidendatensatz. a) ML Baum erstellt aus den kompletten und b) dem RASA Datensatz mit PROTML. c) PROTDET Baum des kompletten Datensatzes. d) PROTDET Baum des RASA Datensatzes

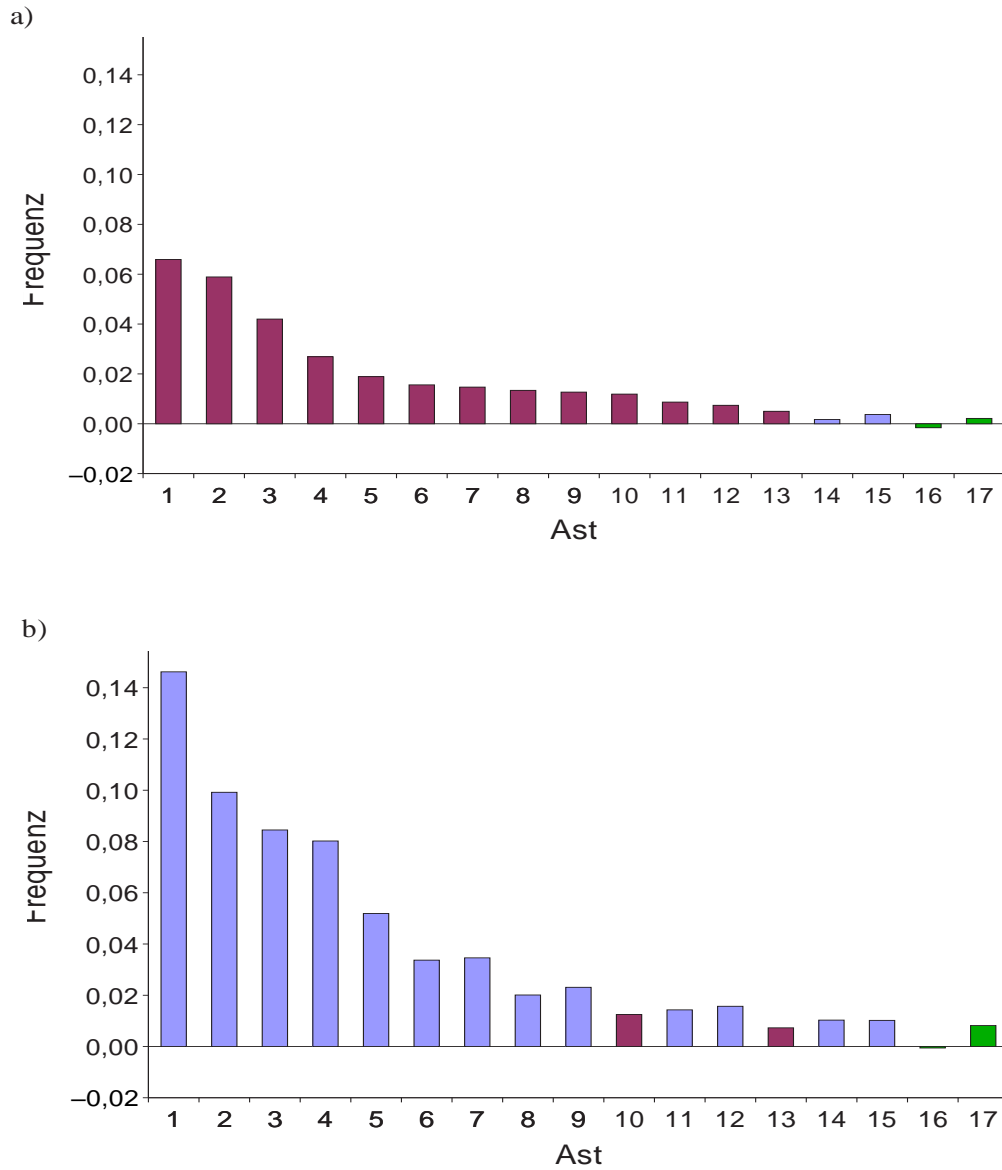


Abbildung 3.12: Die Frequenzen der Splits, die in Tabelle 3.7 aufgeführt sind, wurden mit PROTDET für das gesamte (a) und das RASA Alignment (b) erstellt. Die Splits des gesamten Datensatzes sind in (a) und (b) rot und die des RASA Datensatzes lila markiert. Der grüne Balken der Splits 16 und 17 stehen für die Frequenzen der Splits, die in den ML Topologien nicht aber in den PROTDET Topologien enthalten sind.

Kapitel 4

Diskussion

Eine Gegenüberstellung der Ergebnisse zu bisherigen Forschungsergebnissen gestaltet sich schwierig, weil bisher in der Literatur kaum vergleichbare Untersuchungen existieren.

Bakterienstammbäume basieren in der Literatur zumeist auf Untersuchungen einzelner Proteine vorwiegend 16s rRNA. Lediglich in Wächtershäuser (1998) ist eine Gegenüberstellung von Bakteriengenomen zu finden. Dieser versucht allerdings anhand der Anordnung und Existenz der Gene auf das Genom einer möglichen Urlebensform zu schließen.

Die Untersuchungen der Endosymbiontenhypothese anhand der Gegenüberstellung aller homologen plastidären Proteine und entsprechender Homologer des Cyanobakteriums *Synechocystis sp.* stellen eine Fortführung der vorwiegend institutsinternen Forschungsergebnisse in Martin *et al.* (1998) und Stoebe *et al.* (1999) dar.

Die Kombination der RASA Methode mit den verwendeten phylogenetischen Rekonstruktionsmodellen wurde noch nie in dieser Form durchgeführt. Die RASA Methode für sich alleine diente bisher zum Nachweis von phylogenetischen Signalen, aber nicht als Hilfsmethode phylogenetisch relevante Informationen innerhalb großer Datenmengen zu lokalisieren. Eine andere Methode, die versucht systematisch eine Datenmenge auf ein Minimum an phylogenetisch informativen Daten zu reduzieren, wurde bisher nicht publiziert.

In diesem Kapitel sollen zunächst die Ergebnisse basierend auf der Interaktion von Daten und der Kombination der verschiedenen Modelle diskutiert werden, bevor in den Abschnitten 4.2 und 4.3 eine Gegenüberstellung der erstellten Abstammungsbäume zu bisher bekannten Topologien folgt.

4.1 Die angewandten phylogenetischen Methoden im Vergleich

4.1.1 Die RASA Untersuchungen

Die Ergebnisse der RASA Untersuchungen der gesamten Datensätze sowie der einzelnen Variabilitätsklassen zeigten, daß das RASA Kriterium (vgl. Abschnitt 2.1.4, Lyons-Weiler and Hoelzer (1996), Lyons-Weiler and Hoelzer (1997), Lyons-Weiler and Hoelzer (1999)) ein geeigneter Indikator für das Vorhandensein phylogenetischer Signale darstellt.

Zunächst verdeutlichte die Analyse des Archae-/ Eubakteriendatensatzes, daß sich die Proteine des Datensatzes der Archaeobakterien sehr stark ähneln und sich deutlich von den Eubakterien unterscheiden.

Bei der Untersuchung des RASA Datensatzes wies sowohl der gesamte Datensatz wie auch hauptsächlich niedrige Variabilitätsgruppen das Vorhandensein von phylogenetischen Signalen auf. Dies bestätigte die Hypothese, daß gerade die niedrigen Variabilitätsklassen wichtige phylogenetische Informationen beinhalten.

Zu einem ähnlichem Resultat führten auch die RASA Analysen des Plastidendatensatzes. Auch hier war ein phylogenetisches Signal vermehrt in den niedrigen Variabilitätsklassen zu finden.

Die Auswahl der Variabilitätsklassen des RASA Alignments unterstützt zudem das RASA Kriterium als Indikator für phylogenetische Informationen. Beim Eubakterien Datensatz erstellte PROTML für den gesamten sowie für den RASA Datensatz identische Topologien. Beim Plastidendatensatz waren diese zwei Topologien fast identisch. Hierbei umfaßten die RASA Daten in beiden Fällen weniger als die Hälfte der Daten des gesamten Alignments.

Der Ansatz anhand der Variabilitätsklasseneinteilung und der RASA Analyse informative Daten in einem Alignment zu lokalisieren, stellt zumindest für die PROTML Analyse eine geeignete Möglichkeit dar. Allerdings sprechen die mit PROTDET erstellten Topologien nicht für dieses Vorgehen. Wie im folgenden Abschnitt beschrieben, führt die Reduzierung der zu untersuchenden Daten, bei PROTDET zu fehlerhaften Ergebnissen.

4.1.2 PROTML contra PROTDET

Bei den im vorherigen Kapitel beschriebenen Topologien der Eubakterien- und den Plastidalignments sind zwischen den PROTML und PROTDET Bäumen zum Teil markante Differenzen zu erkennen. Zudem unterscheiden sich die PROTDET Topologien des gesamten und des RASA Datensatzes sowohl bei den Eubakterien als auch bei den Plastidendaten. Besonders auffällig sind

diese Unterschiede bei den Eubakterienbäumen zu beobachten.

Der Hauptgrund dieser starken Differenzen liegt in der Berechnung der *logdet* Distanz. Im Gegensatz zu den von PROTML verwendeten Übergangswahrscheinlichkeiten hängt die *logdet* Distanz zwischen zwei Sequenzen maßgeblich von dem Anteil der übereinstimmenden Aminosäurepositionen ab. Wie in Abschnitt 2.1.3 beschrieben, folgt aus einem hohen Anteil identischer Aminosäuren eine geringe Distanz. Sind in einem Datensatz nur wenig Positionen mit niedriger Variabilität vorhanden, so ist der Anteil identischer Aminosäurepaare wesentlich geringer und folglich erhöhen sich die Distanzwerte. Dies stellt gerade beim Auffinden tief liegender innerer Äste ein Problem dar.

Der Anteil der Positionen mit niedriger Variabilität ist beim Bakteriendatensatz wesentlich geringer als beim Plastidendatensatz. Die RASA Datensätze enthalten keine konservierten Positionen der Variabilitätsklasse 1. Im Falle des Bakteriendatensatzes beinhaltet der RASA Datensatz nicht die Variabilitätsklassen 2 und 3. Im Falle des Plastidendatensatzes fehlt die Variabilitätsklasse 3. Dies führt gerade bei tieferliegenden Ästen zu Problemen. Solche Äste leiten sich aus Sequenzen her, die vorwiegend in den niedrigen Variabilitätsklassen übereinstimmen.

Die Unterschiede zwischen den RASA Topologien der zwei untersuchten Datensätze liegen neben den tiefer liegenden Ästen in solchen Ästen, die im Gesamtdatensatz nur geringe Frequenzen aufwiesen. Dies läßt darauf schließen, daß die Informationen für die Äste hauptsächlich in den niedrigen Variabilitätsklassen zu finden sind, die das RASA Alignment nicht beinhaltet.

Bei den ML Topologien sind die Unterschiede zwischen den Topologien des gesamten und des RASA Datensatzes lediglich in den LBP Werten zu sehen. Hier lassen sich aber ebenfalls ähnliche Tendenzen wie bei den PROTDET Analysen feststellen. Bei den Ästen, deren LBP Werte in der ML Topologie des RASA Datensatzes deutlich unter denen des Gesamtdatensatzes lagen, handelte es sich wiederum um tiefer liegende Äste oder solche, die auch beim Gesamtdatensatz niedrige LBP Werte aufwiesen.

4.2 Untersuchung der Archae- und Eubakterien

Die PROTDET Topologien in Abbildung 3.6 weisen gegenüber der 16s rRNA Phylogenie, die in Abbildung 4.1 nochmals wiedergegeben ist, einige Differenzen auf. Der Grund dürfte in den bereits geschilderten Problemen der PROTDET Analyse des Eubakteriendatensatzes liegen.

Der gesamte sowie der RASA Datensatz weisen bezüglich der Klassifizierung der Bakterienarten für die PROTDET Methode fast identische Informationen wie die der 16s rRNA Analyse auf. Ausnahmen bilden zum einen die Anord-

nung der Purpurbakterien. Die δ/ϵ -Gruppe ist in beiden Topologien nicht zu den übrigen Purpurbakterien benachbart. Zum anderen ist *Mycobacterium* nicht mit den übrigen Gram-positiven Bakterien an einen Ast zu finden.

Die Positionen von *Deinococcus* und *Synechocystis* stellen in allen Topologien nicht sichere Äste dar. In den ML Bäume steht *Synechocystis* näher zu den Gram-positiven *Mycobacterium*. Im Gegensatz dazu erkennen die NJ und die PROTDET Methode eine nähere Verwandtschaft von *Deinococcus* zu *Mycobacterium*. Eine eindeutige Platzierung ist auch anhand der errechneten LBP und Frequenzwerte nicht zu erkennen, liegen diese Werte doch in sehr niedrigen, schwachen Bereichen.

Die ausgewerteten Daten enthalten neben den aus Untersuchungen bereits bekannten Ähnlichkeiten zwischen *Deinococcus* und Cyanobakterien fast ebenso starke phylogenetische Informationen über die enge evolutionäre Beziehung von *Deinococcus* zu Gram-positiven Bakterien.

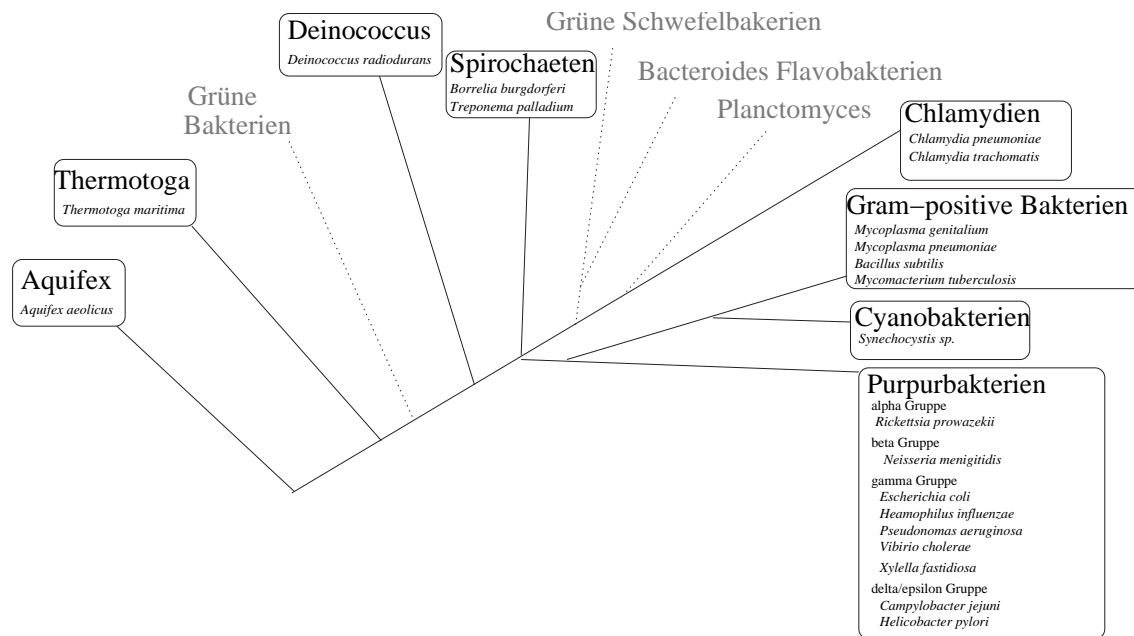


Abbildung 4.1: Zuordnung der 21 Eubakterien zum 16s rRNA Baum

4.3 Die Untersuchung des Plastidendatensatzes

Die Topologien in Abbildung 3.11 weisen für beide Methoden größere Ähnlichkeit mit dem bisher aufgestellten Stammbaum auf, der in Abbildung 4.2 nochmals wiedergegeben ist.

Sowohl die PROTML als auch die PROTDET Methode geben für beide Datensätze die Teilung zwischen den „grünen“ und „roten“ Ästen wieder.

Die evolutionären Beziehungen zwischen den Landpflanzen sind ebenfalls eindeutig erkennbar. Aus den vier Topologien geht außerdem *Mesostigma* als „Urpflanze“ hervor.

Der Ursprung ist in allen Bäumen durch die Cyanobakterie *Synechocystis* und die Glaucocyste *Cyanophora* gut zu erkennen. Allerdings enthält der RASA Datensatz bezüglich der PROTML Methode nicht ausreichend Informationen, um die direkte endosymbiontische Beziehung zwischen diesen Organismen darzustellen.

Im „roten“ Ast zeigen alle Topologien die Ähnlichkeit der Rotalge *Porphyra* und der Cryptophyte *Guillardia*. Dieser Ast unterlegt, daß die Chloroplasten von *Guillardia* durch Aufnahme einer Rotalge in einer sekundären Endosymbiose entstanden sind. Eine ähnliche Beziehung lassen auch die ML Bäume zwischen *Cyanidium* und *Odontella* zu. Die PROTDET Bäume können dies allerdings nicht eindeutig unterlegen.

Die sekundär endosymbiontische Beziehung von *Euglena* zu den Grünalgen spiegeln sich ebenfalls nur die ML Topolgien wider.

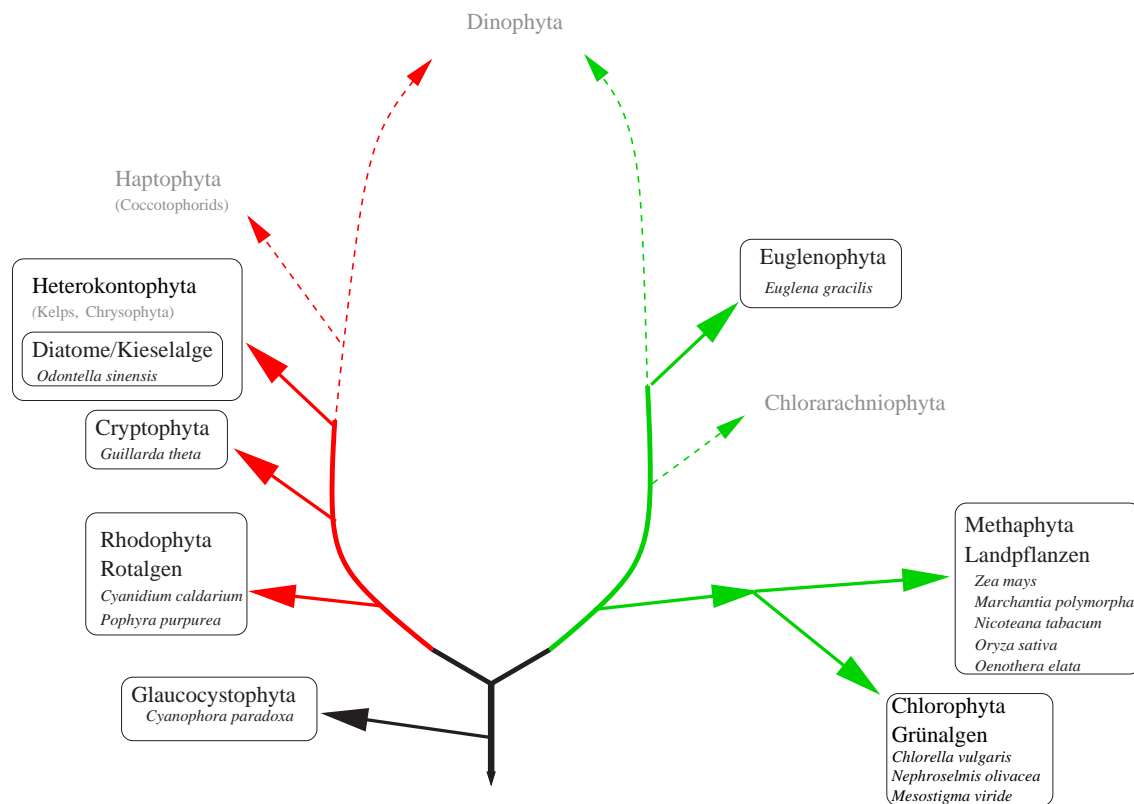


Abbildung 4.2: Zuordnung der 9 Algenarten und 6 Landpflanzen zum Abstammungsbaum der Landpflanzen und Algenarten

4.4 Bewertung

Die erstellten Topologien zeigten, daß die Datensätze ausreichend phylogenetische Informationen enthalten, um weitgehend aufgelöste Phylogenien zu erstellen. Die Vorgehensweise durch die RASA Methode und der Sortierung nach Variabilitätsklassen den Datensatz auf phylogenetisch informative Abschnitte zu reduzieren erwies sich für die mit PROTML erstellten Bäume als erfolgversprechend.

Die mit dem Programm PROTDET erstellten Topologien wiesen, wie beschrieben, Differenzen zu den anerkannten Topologien auf. Bei der Untersuchung der Bakterienevolution wird dies besonders deutlich. Auch die Reduzierung der Daten zum RASA Alignment wirkt sich auf die PROTDET Analysen nicht vorteilhaft aus. Der entscheidende Faktor bei der PROTML Analyse ist die Berechnung der Distanz mit der *logdet* Methode. Hierbei stellt der Anteil identischer Aminosäuren zweier Aminosäuren gegenüber dem Anteil der unterschiedlichen Aminosäuren den entscheidenden Faktor dar. Übergangswahrscheinlichkeiten mit der eine Aminosäure durch eine andere ersetzt wird, bildet die Grundlage der PROTML Analysen. Auf die Berechnung der *logdet* Distanzen haben diese allerdings keinen Einfluß. Bei Datensätzen deren Anteil an konservierten Positionen nur sehr gering ist, wie die RASA Alignments oder der Bakteriendatensatz führt dies zu einem erheblichen Problem beim Bestimmen der Distanzen. Durch Einbeziehung der von Übergangswahrscheinlichkeiten und eine effizientere Behandlung von Sonderfällen (z.B. Determinante hat den Wert Null) könnte die Berechnung der *logdet* Distanzen und somit der PROTDET Bäume sicherlich deutlich stabilisieren.

Allgemein läßt sich sagen, daß bei Alignments mit einem geringen Anteil konservierter Positionen oder bei kleinen Datensätzen die PROTML Analysen verlässlichere Ergebnisse liefern. Ist die Datenmenge allerdings groß und enthält einen hohen Anteil konservierter Positionen so bietet sicherlich die PROTDET Analyse eine effiziente Alternative.

Kapitel 5

Zusammenfassung und Ausblick

In dieser Arbeit wurden mit Hilfe unterschiedlicher phylogenetischer Methoden drei Datensätze analysiert, die sich jeweils aus homologen Proteinen der untersuchten Organismen zusammensetzten. Ein Datensatz beinhaltete homologe Proteine von 21 Eubakterien, ein weiterer zusätzlich die Gene von 5 Archaeobakterien. Der dritte umfaßte homologe Proteine der Cyanobakterie *Synechocystis sp.* und der entsprechenden plastidär kodierten Proteine von 6 Landpflanzen sowie 9 Algenarten. Ziel war es zu überprüfen, ob diese Daten, die die gesamten Genome berücksichtigten, ausreichend bzw. neue evolutionäre Informationen beinhalteten, um eine nahezu vollständige phylogenetische Auflösung zu erzielen. Weiterhin sollten Möglichkeiten untersucht werden, im Datensatz phylogenetisch informative Abschnitte zu lokalisieren, um die Datensätze auf diese reduzieren zu können.

Zunächst wurden zu den einzelnen Proteinen Alignments erstellt, aus denen das Programm PROTML (Adachi and Hasegawa (1996)) phylogenetische Bäume rekonstruierte. Die Alignments der Proteine, deren Abstammungsbäume nicht auf gesonderte evolutionären Prozesse wie lateralen Gentransfer schließen ließen, ergaben durch Konkatenierung die zu untersuchenden Datensätze.

Mit dem Programm SORTAL (Abschnitt 2.1.5.2) ließen sich die Alignments nach Variabilität sortieren. Mit der RASA Analyse (Lyons-Weiler and Hoelzer (1996), Lyons-Weiler and Hoelzer (1997), Lyons-Weiler and Hoelzer (1999), Abschnitt 2.1.4) konnten nun in den sortierten Datensätzen die Variabilitätsklassen identifiziert werden, die ein phylogenetisches Signal enthalten. Für den Datensatz der 5 Archae- und 21 Eubakterien ergab sich wider Erwarten, daß gerade die niedrigen Variabilitätsklassen auf kein phylogenetisches Signal hindeuteten. Dieses Verhalten ließ sich durch die starke Ähnlichkeit der Archaeobakterien untereinander und der erheblichen Unterschiede zu den Eubakterien begründen. Die weiteren Untersuchungsschritte erfolgten deshalb nur für die zwei anderen Datensätze.

Anhand der RASA Ergebnisse wurden zu den zwei übrigen Datensätzen RASA Alignments (Abschnitt 2.1.5.4) erstellt, die nur aus den Variabilitätsgruppen bestanden, die laut der RASA Methode (Lyons-Weiler and Hoelzer (1996), Lyons-Weiler and Hoelzer (1997), Lyons-Weiler and Hoelzer (1999)) ein phylogenetisches Signal aufwiesen. Zu den vier Datensätzen rekonstruierten die Programme PROTDET (Hendy and Penny (1993), Lockhart *et al.* (1994), Penny and Hendy (2001)) und PROTML (Adachi and Hasegawa (1996)) jeweils einen phylogenetischen Baum.

Im Vergleich zu bekannten Topologien zeigte sich, daß sich die mit PROTML erstellten Bäume kaum von diesen unterschieden. Bei PROTDET Bäumen waren allerdings wesentliche Differenzen besonders beim Bakteriendatensatz und bei den RASA Datensätzen zu erkennen. Dies ließ sich durch die geringe Anzahl an konservierten Positionen des Bakterienalignments bzw. durch den Wegfall dieser in den RASA Alignments erklären.

Aufgrund dieser Ergebnisse konnten die folgenden Rückschlüsse gezogen werden:

- In den Datensätzen sind auch nach Reduzierung auf den RASA Datensatz ausreichend phylogenetische Informationen bezüglich der Rekonstruktion mit dem Programm PROTML enthalten. Für PROTDET gilt diese Aussage nur bedingt.
- Die PROTDET Methode in dieser Form eignet sich nur zur phylogenetischen Rekonstruktion, wenn die Datensätze groß sind und einen hohen Anteil an konservierten Abschnitten besitzen. Durch Erweiterungen und bessere Behandlung von Sonderfällen in der Berechnung ließe sich diese Methode stabilisieren.
- Auch wenn die Bäume nur wenig neue Informationen lieferten, stellt sich die Untersuchung ganzer Genome als künftige Herausforderung dar. Die Kombination aus Sortieren der Alignments mit SORTAL und Lokalisieren der phylogenetisch informativen Abschnitte mit RASA bieten einen guten Ansatzpunkt, die zu untersuchende Datenmenge zu reduzieren

Auch wenn die Ergebnisse vielversprechende Ansätze in der phylogenetischen Forschung bieten, bleibt bei diesem Vorgehen der Zeitfaktor beim Erstellen der Datensätze zu berücksichtigen. Um die Untersuchung ganzer Genome zu etablieren, müßten Datenbankprogramme erstellt werden, die es auf schnellem Wege ermöglichen, homologe Proteine lokalisieren und diese zu Datensätzen zusammenstellen zu können. Durch die kontinuierlich steigenden Rechnerleistungen wird die Zeitfrage bei der Rekonstruktion phylogenetischer Bäume immer mehr in den Hintergrund treten. Programme wie PROTDET werden in ausgereifterer Form Alternativen zu den gebräuchlichen Distanzmatrix- oder ML Methoden bieten.

Anhang A

Benutzerhandbuch

In diesem Kapitel sollen kurz die Bedienung der für diese Arbeit implementierten Programme beschrieben werden. Die Programme sind für UNIX/LINUX Plattformen entwickelt worden. Das Programm PBL verwendet zur Rekonstruktion der phylogenetischen Bäume die PHYLIP Programme (Felsenstein (1993)) PROTPARS, PROTDIST, NEIGHBOR, SEQBOOT sowie CONSENSE und die MOLPHY Programme (Adachi and Hasegawa (1996)) NJDIST und PROTML. Um PBL nutzen zu können, müssen diese Programme für den Benutzer ausführbar sein.

A.1 Eingabeformat des Alignments

Für die in diesem Abschnitt vorgestellten Programme wird als Eingabedatei ein Alignment im sogenannten „PHYLIP-Format“ benötigt. Hierbei steht in der ersten Zeile die Anzahl der Sequenzen und die Sequenzlänge. In den weiteren Zeilen folgen die Sequenzen, die blockweise angegeben sind.

Die Zeilenanzahl eines Blockes entspricht der Sequenzanzahl. Die Spalten eines Blockes werden in 6 Gruppen zu je 10 Zeichen eingeteilt, die jeweils durch ein Leerzeichen voneinander getrennt sind.

Die erste Gruppe im ersten Block enthält die Namen der Sequenzen, die gegebenenfalls durch Leerzeichen auf 10 Zeichen ergänzt werden. In den weiteren Blöcken enthält die erste Gruppe nur Leerzeichen.

Die Gruppen 2-6 geben die durch Großbuchstaben spezifizierten Nukleotide/Aminosäuren der Sequenzen wieder. Ein Block enthält somit einen Abschnitt von 50 Positionen des Alignments. Falls die Positionsanzahl nicht durch 50 teilbar ist, enthält der letzte Block nur noch die restlichen *Positionsanzahl MOD 50* Positionen.

A.2 Das Programm SORTAL

Das Programm SORTAL wird mit

```
sortal "dateiname"
```

oder falls das zu untersuchenden Alignment in der Datei „infile.dat“ abgespeichert wurde durch

```
sortal
```

aufgerufen.

Die Datei „outfile“ beinhaltet das sortierte Alignment. In der Datei „posvar.dat“ sind Informationen über die Variabilität der Positionen des unsortierten und sortierten Alignment zu finden.

A.3 Das Programm PBL

Das Programm wird mit

```
PBL "dateiname"
```

oder nur mit

```
PBL
```

aufgerufen. Im zweiten Fall sollte das zu untersuchende Alignment unter dem Dateinamen „infile.dat“ verfügbar sein.

```
The alignment has 20 OTUs and 413 positions
*****
Create new consensus trees? (y/n)
```

Sollen neue phylogenetische Bäume zu dem Datensatz erstellt werden, dann ist die Frage mit „y“ beantworten. Falls aber schon mit PBL Bäume erstellt und die Informationen in der dazugehörigen Datei „edges“ zu finden sind, kann die Frage mit „n“ beantwortet werden (s.u.).

Durch die Eingabe von „y“ gelangt man zu den folgenden 6 Menüepunkten.

1. Sort Alignment (no)
2. Method (protpars/nj/protml) (nj-kimura with 1 samples)
3. Bootstrapping (no)
4. Reduce/Window (reduce by 413 positions)
5. Start
6. Exit

Your Choice:

- Your Choice : 1

Do you want to sort the alignment by positional variability? (y/n)

Es besteht die Möglichkeit, den zu untersuchenden Datensatz mit SORTAL zunächst zu sortieren. Bei unsortierten Datensätzen lassen sich weiter unten aufgeführte Optionen nicht sinnvoll einsetzen.

- Your Choice : 2

1. Protpars
2. Neighbor-Joining
3. Maximum Likelihood

Mit dem Menüpunkt zwei läßt sich das Programm einstellen, mit der die phylogenetischen Bäume erstellt werden sollen. Zur Auswahl stehen die zwei PHYLIP Programme PROTPARS (Parsimony Methode) und NEIGHBOR (Distanzmatrix Methode). Mit dem Programm PROTML werden die Bäume mit dem Maximum Likelihood Programm des MOL-PHY Packetes rekonstruiert.

Choose number of option

1. Kimura Formula
2. Dayhoff

Im Falle der Neighbor-Joining Methode, kann zwischen zwei Methoden zur Berechnung der Distanzmatrix ausgewählt werden. Die Kimura Formel ist eine recht schnelle Methode die Distanzen zwischen den Sequenzpaaren abzuschätzen und berechnet sich ledig aus dem paarweisem Vergleich der Sequenzen. Sie kann aber auch zu Fehlabschätzungen führen. Bei genaueren Berechnungen sollte der zeitaufwendigeren Methode nach Dayhoff der Vorzug gegeben werden. Anhand der Dayhoff

Martix, die die Substitutionswahrscheinlichkeiten zwischen den 20 Aminosäuren wiedergibt, und dem paarweisem Sequenzvergleich ergeben sich die Dayhoff Distanzen für alle Sequenzpaare.

Die Berechnungen der Distanzen wird durch das Programm PROTDIST realisiert, das ebenfalls im PHYLIP Packet enthalten ist.

- Your Choice : 3

Bootstrapping (y/n)?

Please enter number of bootstrap samples to run

Die Option Bootstrapping läßt sich nur in Verbindung mit einem der zwei PHYLIP Programme PROTPARS und NEIGHBOR anwenden. Bei der Bootstrap Option werden aus dem zu untersuchenden Alignment eine vom Anwender angegebene Anzahl von Alignments mit der gleichen Positionsanzahl erstellt. Die Positionen dieser Alignments entsprechen Positionen des zu untersuchenden Alignments, die zufällig ausgewählt wurden. Positionen des Ausgangsalignments können dabei auch wiederholt ausgewählt werden. Dieses Vorgehen, nach dem Prinzip „Ziehen mit zurücklegen“, wird durch das PHYLIP Programm SEQBOOT ausgeführt.

Bei diesem Vorgehen wird aus jedem dieser so erstellten Alignments ein Baum rekonstruiert. Aus allen diesen Bäumen erstellt das Programm CONSENSE einen Consensus Baum, der sich aus den sich nicht widersprechenden Ästen zusammensetzt, die in den erstellten Bäumen am häufigsten enthalten sind.

- Your Choice : 4

(V) Remove the position-group with the highest variability
(R) Reduce alignment
(G) Variability Groups
(W) Window

Durch den 4.Menuepunkt läßt sich die Methode bestimmen, mit der aus einem sortiertem Alignments gezielt Teilalignments zur phylogenetischen Rekonstruktion herangezogen werden können.

Die Option (V) ermöglicht es, das zu untersuchenden Alignment um jeweils die Positionen der höchsten Variabilitätsklasse zu reduzieren und daraus phylogenetische Bäume zu erstellen. Durch die Option (G) lassen sich die Positionen einer Variabilitätsklasse einzeln untersuchen. Bei diesen beiden Optionen lassen sich die Variabilitätsklassen definieren, bei der die Untersuchungen beginnen und enden soll.

```
Start at Variability group
Stop at Variability group
```

Die Option (R) ermöglicht es, das Alignment vor jedem Untersuchungsschritt um eine konstante Anzahl der variabelsten Positionen zu reduzieren.

```
By how many positions to reduce the alignment each time?
```

Bei der Option (W) wird für jeden Untersuchungsschritt ein Ausschnitt, ein sogenanntes Fenster, mit jeweils einer konstanten Positionsanzahl *ws* ausgewählt. Dieses Fenster wird in jedem Schritt um die konstante Positionsanzahl *shift* beginnend von den hintersten bis zu den vorderen Positionen verschoben. Das bedeutet für ein mit SORTAL sortiertes Alignment von den variabelsten bis zu den konserviertesten Positionen. Die Werte *ws* und *shift* sind vom Anwender zu definieren.

```
Please enter windowlength :
Please enter step size ( window shift )
```

- Your Choice : 5

```
Start PBL
```

Mit dem Menüepunkt 5 startet man die Auswertungen mit PBL.

- Your Choice : 6

Quit PBL

Durch den Menüpunkt 6 wird das Programm abgebrochen.

Nach Beendigung der Auswertungen, ermöglicht PBL dem Benutzer sich BP-Werte für Äste angeben zu lassen, die er selber definiert. Wird am Anfang des Programms die Frage

Create new consensus trees? (y/n)

mit 'n' beantwortet, gelangt man gleich an diesen Punkt des Programms. Es erscheint nun ein weiteres Menü.

Options

(E) Enter edge

(Q) Quit

Das Programm kann nun mit 'q' beendet werden. Durch Eingabe von 'e' bietet sich die Möglichkeit bestimmte Äste zu definieren und sich dazugehörigen Bootstrap Werte in die Datei „edge.dat“ ausgeben zu lassen. Dies geschieht durch Angabe aller OTUs, die auf der einen Astseite zu finden sind.

Enter edge, please

List of all OTUs

1. OTU1
2. OTU2
3. OTU3
4. OTU4
5. OTU5
6. OTU6
7. OTU7
8. OTU8
9. OTU9
10. OTU10
11. OTU11
12. OTU12
13. OTU13
14. OTU14
15. OTU15
16. OTU16
17. OTU17
18. OTU18

19. OTU19
20. OTU20

Please enter OTU's numbers :

Nun gibt man die Nummern der OTUs ein, die sich auf einer Seite des Astes befinden, dessen BP-Werte man wissen möchte. Die Nummern der OTUs werden durch ein Komma getrennt. Hinter der letzten OTU-Nummer muß ein '.' stehen.

Durch die Eingabe von

13, 14, 18.

werden die BP-Werte für den Ast in 'edge.dat' vermerkt, an dessen einem Ende sich die OTUs 13,14,18 befinden. Äquivalent könnte man auch

1,2,3,4,5,6,7,8,9,10,11,12,15,,16,17,19,20.

List of all OTUs

- ```

1. OTU1
2. OTU2
3. OTU3
4. OTU4
5. OTU5
6. OTU6
7. OTU7
8. OTU8
9. OTU9
10. OTU10
11. OTU11
12. OTU12
13. OTU13 13. OTU13
14. OTU14 14. OTU14
15. OTU15
16. OTU16
17. OTU17
18. OTU18 18. OTU18
19. OTU19
20. OTU20

```

Press any key to continue

Es wird nun eine Liste der OTUs ausgegeben, bei der die eingegebenen OTU Nummern markiert sind. Durch einen beliebigen Tastendruck gelangt man wieder zu dem Menue:

```
Options
(E) Enter edge
(Q) Quit
```

In der Datei 'edge.dat' werden die BP-Werte für die Äste eingetragen, die der Benutzer definiert hat.

```
1 : OTU13 OTU14 OTU18
413 100
313 0
213 0
113 0
13 0
```

Wurde eine der Option (R) oder (V) verwendet, so entspricht die erste Zahlenangabe der Anzahl der Positionen des Alignments. Bei den anderen Optionen wird die mittlere Positionsnummer des jeweiligen Alignment-Ausschnittes ('window') angegeben. In diesem Fall wurde die 'Reduce-Option' um 100 Positionen durchgeführt. Der oben definierte Ast konnte nur aus dem Alignment der Länge 413 rekonstruiert werden.

## A.4 Das Programm RASA

Das RASA Programm wird mit

```
unixRASA "dateiname"
```

oder falls das zu untersuchenden Alignment in der Datei „infile.dat“ abgespeichert wurde durch

```
unixRASA
```

Für die Untersuchung muß lediglich die Anzahl der Permutationen zur Berechnung der *tRasa* angegeben werden.

Enter Number of Permutations (max. 1000 )

Nach der Bildschirmausgabe der berechneten Werte wird das Programm mit der Meldung

Finished Rasa

beendet.

Die berechneten *Eij* und *Rasa* Werte sind in den Dateien „vrasaXX“ für die einzelnen Variabilitätsgruppen und für den gesamten Datensatz in der Datei „TRasa“ wiedergegeben. Die *tRasa* Werte bezüglich der einzelnen Variabilitätsgruppen sind in der Datei „tRasa.p“ zu finden.

Die Werte in diesen Dateien sind so angegeben, daß sie sich aus ihnen mit dem Programm GNUPLOT Diagramme erstellen lassen. Das RASA Programm stellt mit der Datei „draw“ eine Eingabe Datei zur Verfügung, um mit dem Aufruf

`gnuplot<draw`

Diagramme im Postscript Format *ps* und im XFig Format *fig* erstellen. Falls das XFig Programm vorhanden ist lassen sich die *fig* Dateien betrachten und in ein gewünschtes Format exportieren. Zum Export der Dateien in das *gif* Format kann die Datei „figTOgif“ ausgeführt werden. Die Ergebnisdiagramme und berechneten Werte lassen sich in der Datei „rasa.htm“ in einem beliebigen Browser betrachten. Das RASA Alignment wird in der Datei „outfile.phy“ gespeichert.

## Anhang B

# Implementierung der Programme

Die in dem Rahmen dieser Arbeit implementierten oder weiterentwickelten Programme wurden unter dem Betriebssystem UNIX/LINUX in C/C++ erstellt. Es wurde dabei der GNU C++ Compiler verwendet (Version 2.95.2).

Die Datenstruktur, die in Abbildung B.1 dargestellt ist, verwaltet die für die Programme SORTAL, PBL und RASA relevanten Informationen des zu untersuchenden Alignments.

Die Daten des Alignments werden zu Beginn dieser Programme positionsweise in eine Queue eingelesen. In jedem Element dieser Queue werden neben einem Verweis auf das folgende Queue Element, die Positionsnummer, die Variabilität, die prozentuale Verteilung der Aminosäuren und „gaps“ sowie die Buchstaben der Aminosäuren und „gaps“ einer Position, ebenfalls als Queue, abgelegt.

Für das Programm SORTAL wird für jedes Element zusätzlich eine Priorität, wie Abschnitt 2.1.5.2 beschrieben berechnet, anhand der das Element in die Queue einsortiert wird. Durch die Ausgabe der Positionen der sortierten Queue erhält man das sortierte Alignment.

Im Programm PBL ermöglicht diese Datenstruktur aus dem sortierten Datensatz einfach und gezielt die möglichen Teildatensätze zur phylogenetischen Untersuchung auszuwählen.

Die Datenstruktur ermöglicht bei der RASA Untersuchung die Berechnung der *tRasa* Werte bezüglich der einzelnen Variabilitätsklassen, selbst wenn das Alignment vorher nicht mit SORTAL sortiert wurde.

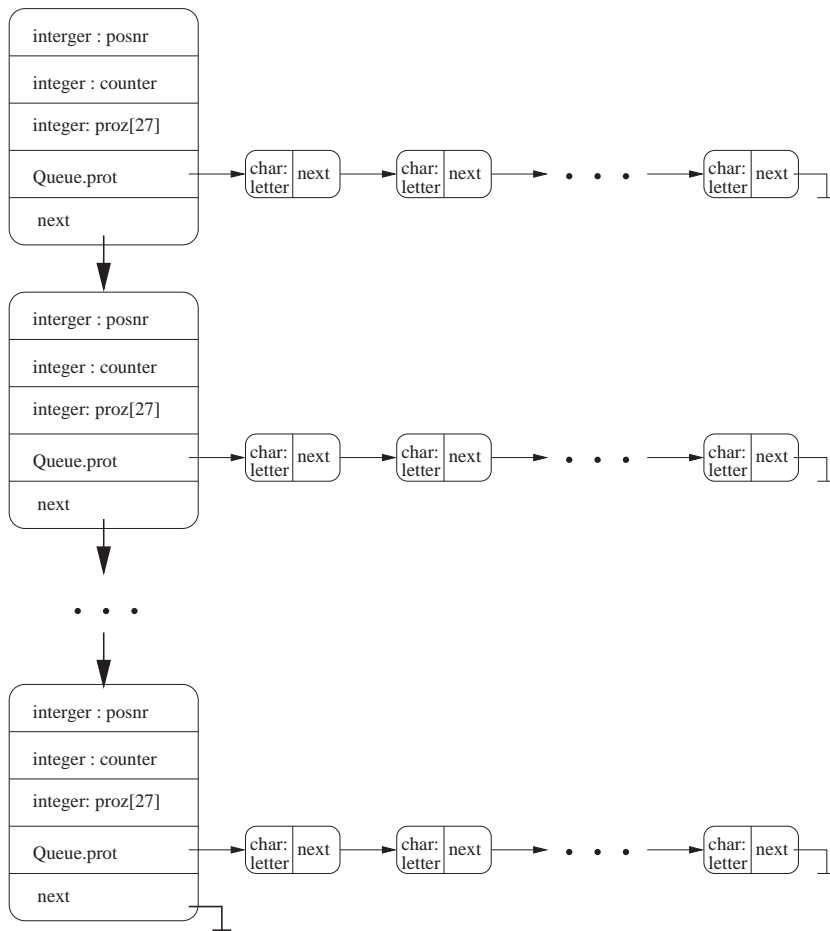


Abbildung B.1: Die Informationen eines Alignments werden in einer Queue abgelegt. Jedes Queue Element beinhaltet die Informationen einer Position des Datensatzes: Positionsnummer (posnr) und Anzahl unterschiedlicher Aminosäuren einer Position (counter) werden durch ganzzahlige positive Zahlen (integer) wiedergegeben. In dem integer Feld *proz* sind die prozentualen Häufigkeiten der durch Buchstaben repräsentierten Aminosäuren und den „gaps“, dargestellt durch ‘-’, festgehalten. Die Buchstaben, die die Aminosäuren und „gaps“ einer Position darstellen, sind der Queue prot gespeichert.

# Anhang C

## Der Eubakteriendatensatz

01aae ----MARE---VPIEKLNRNIGIVAHIDAGKTTTTERILYTTGKTYKIGEVH--EGA-ATM  
04bb -----MD-----YNKLRNIGISAHIDSGKTTTLTERILFYCNKIHAHEVKGKDGVGATM  
05bs ----MAREFS---LEKTRNIGIMAHIDAGKTTTTERILFYTGRIHKIGETH--EGA-SQM  
06cj ----MSRS---TPLKKVRNIGIAAHIDAGKTTTTERILFFTGM SHKIGEVH--DGA-ATM  
07cp ----MSNQE---FDLSAIRNIGIMAHIDAGKTTTTERILFYAGRTHKIGEVH--EGG-ATM  
08ct ---MSDQE---FGLDAIRNIGIMAHIDAGKTTTTERILFYAGRTHKIGEVH--EGG-ATM  
09der ----MTITKSQQYLTHFRNIGIAAHIDAGKTTTTERILFFTGRIRNLGETH--EGA-SQM  
10ec ----MART---TPIARYRNIGISAHIDAGKTTTTERILFYTG VNHKIGEVH--DGA-ATM  
11hi ----MART---TPIERYRNIGISAHIDAGKTTTTERILFYTG VSHKIGEVH--DGA-ATM  
12hp ----MARK---TPLNRIRNIGIAAHIDAGKTTTTERILFYTG VSHKIGEVH--DGA-ATM  
13mg ----MSRTVD---LKNFRNFGIMAHIDAGKTTTTERILFHSGRHKIGETH--DGE-SVM  
15mp ----MARTVD---LINFRNFGIMAHIDAGKTTTTERILFHSGRHKIGETH--DGE-SVM  
17mtu ----VAQKDVLTDL SRVRNFGIMAHIDAGKTTTTERILYTTG INYKIGEVH--DGA-ATM  
18nm ----MARK---TPISLYRNIGISAHIDAGKTTTTERILFYTG LTHKLGEVH--DGA-ATT  
19pae ----MART---TPINRYRNIGICAHVDAGKTTTTERVLFYTG VNHKLGEVH--DGA-ATT  
21rp ----MSKI---NKLEQIRNIGICAHIDAGKTTTTERILYTTG KSHKIGEVH--EGG----  
22syn ----MARTVP---LERIRNIGIAAHIDAGKTTTTERILFYSG VVHKIGEVH--EGT-AVT  
23tm MQNVEARY---VDLDKLRNIGIMAHIDAGKTTTTERILYTTG RGHKIFGDVD--EGN-TTT  
24tp ----MSRG-----ISTFRNIGISAHIDSGKTTTLSEILFYCDRIHAHEVRGKDGVGATM  
25vib ----MARK---TPIERYRNIGICAHVDAGKTTTTERILFYTG LSHKIGEVH--DGA-ATM  
26xyf ----MVR A---TPIHRYRNIGIMAHIDAGKTTTTERILFYAG VCHQMGEVH--DGA-AVM

01aae DWMPQEKERGITITVATTACYWTRNG---ERYQINIIDTPGHVDFSVEVVRSMKVLGDIV  
04bb DSMELERERGITIASAATHVEWKD-----FPINIIDTPGHVDFTIEVERSLRVLGAI  
05bs DWMEQEQERGITITSAATTACQWK-----YRVNIIDTPGHVDFTIEVERSLRVLGAV  
06cj DWMEQEQERGITITSAATTCFWK-----DHQINLIDTPGHVDFTIEVERSMRVLGAV  
07cp DWMAQEQERGITITSAATTVFWLGA-----KINIIDTPGHVDFTIEVERSLRVLGAV  
08ct DWMEQEQERGITITSAATTVFWLGA-----KINIIDTPGHVDFTIEVERSLRVLGAV  
09der DWMEQERERGITITAAATTAHWTHTGTG-EDYTVNIIDTPGHVDFTIEVERSMRVLGAV  
10ec DWMEQEQERGITITSAATTAFWSGMAKQYEPHRINIIDTPGHVDFTIEVERSMRVLGAV  
11hi DWMEQEQERGITITSAATTAFWSGMSQQFPQHRINVIDTPGHVDFTIEVERSMRVLGAV  
12hp DWMEQEQERGITITSAATTCFWK-----DHQINLIDTPGHVDFTIEVERSMRVLGAV  
13mg DWMEQEQERGITITSAATSVSWKN-----CSLNLIDTPGHVDFTIEVERSLRVLGAI  
15mp DWMEQEQERGITITSAATSVSWKN-----CSLNLIDTPGHVDFTIEVERSLRVLGAI  
17mtu DWMEQEQERGITITSAATTFWK-----DNQLNIIDTPGHVDFTIEVERNLRVLGAV  
18nm DYMEQEQERGITITSAAVTSYWSGMAKQFPEHRFNIIDTPGHVDFTIEVERSMRVLGAV  
19pae DWMVQEQERGITITSAAVTFWKSGRQYDNYRVNVIDTPGHVDFTIEVERSLRVLGAV  
21rp -----  
22syn DWMAQERERGITITAAAISTDWL G-----HHINIIDTPGHVDFTIEVERSMRVLGVI  
23tm DWMPQEKERGITIQAATTCFW-KG-----YRINIIDTPGHVDFTAEVERALRVLGAI  
24tp DNMELEERERGITIQAASTVQWK G-----HTINVIDTPGHVDFTIEVERSLRVLGAV  
25vib DWMVQEQERGITITSAATTFWRGMEAQFQEHRINIIDTPGHVDFTIEVERSLRVLGAV  
26xyf DWMEQEQERGITITSAATTVFWSGMDKMPQHRFNIIDTPGHVDFTIEVERSLRVLGAV

01aae FIFSAVEGVQPQSEANWRWADRFQVPRIAFVNKMDRLGADFYRVFKEIEEKLTIKPVAIQ  
04bb LVLD SVAGVQSQSI TVDRQLKRYSVPRIAFVNKMDKTGANPYNVKDQLRSKLDLNSVLMQ  
05bs AVLDAQSGVEPQ TETVWRQATTYGVPRIVFVNKMDKIGADFLYSVGTLRDRLQANAHAIQ  
06cj AVFC SVGGVQPQSETVWRQANKYGVPRIVFVNKMDRIGANFYNVEDQIRNRLKANVPVLQ  
07cp AVFDA SVGVEPQSETVWRQADKYGVPRIAFVNKMDRMGADYFAAVESMKEKLGANAPVPH  
08ct AVFDA SVGVEPQSETVWRQANKYGVPRIAFVNKMDRMGANFYGAIESMREKLGANAIPVH  
09der AVFDSQGVQPQSETVWRQADRYGVPRIAFVNKMDKTGASFDLVNDIRERLGAIPAPIQ  
10ec MVYCAVGGVQPQSETVWRQANKYKVPRIAFVNKMDRMGANFLKVNVQIKTRLGANPVPLQ  
11hi MVYCAVGGVQPQSETVWRQANKYEVPRIAFVNKMDRTGANFLRVVEQLKTRLGANAIPLQ  
12hp SVFCSVGGVQPQSETVWRQANKYGVPRIVFVNKMDRIGANFYNVENQIKLRLKANVPVIN  
13mg AVLDAQMGVEPQ TETVWRQASRYEVPRIVFVNKMDKTGANFERSVLSIQRLGVKAVPIQ  
15mp AVLDAQMGVEPQ TETVWRQASRYEVPRIVFVNKMDKTGANFERSVQSIIQRLGVKAVPIQ  
17mtu AVFDGKEGVQPQSE VWRQADKYDVPRICFVNKMDKIGADFYFSVRTMGERLGANAVPIQ  
18nm MVYCAVGGVQPQSETVWRQANKYQVPRIAFVNKMDRQGANFVRVEQMKTRLANPVPPIV  
19pae VVFCGTSGVEPQSETVWRQANKYGVPRIVYVNKMDRQGANFLRVVEQIKKRLGHTPVVPVQ  
21rp -----  
22syn AVFC SVGGVQPQSETVWRQAERYQVPRIAFVNKMDRTGANFVRVCQIQGDRLRANAVPVQ  
23tm RVFDATAGVEPQSETVWRQADKYNVPRIAFVNKMDKVGADFYMAVELTVTKLRANPIPVQ  
24tp LVLC SVAGVQSQSI TVDRQLRRYHVPRISFINKCDRTGANFPKCAQLREKLSLAHLMQ  
25vib VVFCGTSGVEPQSETVWRQADKYGVPRMVFNKMDRAGADFLRVVGQIKHRLGANPVPIQ  
26xyf FVLC AVGGVQPQSETVWRQANKYFVPRMAFVNKMDRTGANFDFKVVGEQLKARLGANVPVMQ

01aae IPLGAEDQFEGVIDLMEMKAIRWEETL--GAKYEVVDIPPEYQEKAQEWREKMIETIVET  
04bb IPIGLEDKHIGVIDLVMKAYYFEGKDG--TEIEKEIPSELLEEAKNREMLDALADF  
05bs LPIGAEDNFEGIIDLVENKAYFYEDDLG--TRSDAKEIPPEYKEQAEELRNSLIEAVCEL  
06cj IPIGAEDNFKGVIDLVTMKALVWEDDTK--PTDYVEKEIPAELEKAEYRTKMIEAVSET  
07cp CPIGSESQFVGMVDLISQKALYFLDDTL--GAKWEEKEISEDKERCAELRANLLEELATI  
08ct CPIGSESQFVGMVDLISQKALYFLEETL--GAKWEEKEIPEDLQEQCATLRMQLLEELATV  
09der YPMGAENDFKGVIDIVRMQAHTFTNDLG--TEIQVGDVPAEYMDKVNEMRQQLIEAAAEV  
10ec LAIGAEEHFTGVVDLVMKAIWNNDADQ--GVTFEYEDIPADMVELANEWHQNLIESAAEA  
11hi LPVGAEEENFTGVVDLVMKAIWNNDADQ--GMTFTYEEVPANMQADCEEWRQNLVEAAAEA  
12hp IPIGAEDTFIGVIDLVQMKAIVWNNETM--GAKYDVEEIPSDLLEKAEYREKLVEAAEQ  
13mg FPIGAENDFNGIIDITTKAYFFDGNKE--ENAEKPIPEQYVDQVEKLYNNLVEEVASL  
15mp LPIGAENDFNGIIDLIEEKVYFFDGGKE--EKAEEKPIPDQFKDQVKQRAHLVEEVANF  
17mtu LPVGAEDFEGVVDLVMNAKVMRGETKLGETYDVEIPADLAEQAEYRTKLLEVVAAES  
18nm IPVGAEDNFSGVVDLLKMKSIWNNEVDK--GTTFTYGDIPAELEVETAEEWRQNMIEAAAEA  
19pae LAIGAEEENFVGVDLVMKAIYWNDDDK--GMTYREEEIPAEKDLAEWRSSMVEAAAEA  
21rp -----  
22syn IPIGSEAEFEGVIDLVRMKAYLYKNDLG--TDIQEVP IPDSVVDKTEEYRLRLVESVAEA  
23tm MPIGSEKDFQGVIDLKMKAIYVWSED--GSVYEERDIPPEELREEAMREEMLEKIAEL  
24tp LPIGLEDRLLEGVIDLISLKALYFEGESG--AHVREAP IPEQYQADVKKYRDELIDAAELF  
25vib LNTGAEEEFKGVIDLKMKAIWNNDADQ--GMSFTYEEIPADMLELAQEWNRHLVEAAAEA  
26xyf VPIGAEDQFEGVIDLLKMKAIHWDAAEQ--GTVFEYRDIPIELVDKASKARAFMVEAAAEA  
-----  
01aae D---DELMEXYLEGQE-ISIDELRKALKATIERKLVPLCGSAFNKNGVQPLLDIVIDY  
04bb N---DELMELHMEGKE-VPIEIIYNAIRTGTLALKLCPVFMGSAYKNKGVLQLLDVATRF  
05bs D---EELMDKYLEGEE-LTIDELKAGIRKGTLNVEFYPLVGSFAFNKNGVQLVLDVAVDY  
06cj S---DELMEXYLEGEE-LSLEEIKTGKAGCLSLSIYVPLCGTAFKNKGVQPLLDVAVVY  
07cp DESNEAFMMKVLDPDSITEDEIHQVMKGVENKINPVLGCTAFKNKGVQQLLVIVKW  
08ct DESNEAFMEKVLNPDSDITEEIHVMKGVIEGKINPVLGSAFNKNGVQQLLDVIVKW  
09der D---EDLMMKFLEGEE-PTQELIAAIRKGTIDKQIFPVLGSAFNKNGVQQLLDVAVDY  
10ec S---EELMEKYLEGEE-LTEAEIKGALRQRLVNEIILVTCGSAFNKNGVQMLDAVIDY  
11hi S---EELMEKYLEGEE-LTEEEIKSALRQRLVLANEILVTCGSAFNKNGVQMLDAVVEY  
12hp D---EALMEKYLEGEE-LSIEEIKKGKAGCLNMSLVPMLCGSFAFNKNGVQTLDAVIDY  
13mg D---DQLMADYLDGKP-IEIDAINKAIRNGVIHCKFFPVLGSAFNKNGIKLLDAVVDV  
15mp D---DQLMADYLDGKE-ISAIDIKRCIRKGVIGCGFFPVLGSAFNKNGIKLLDAVVDV  
17mtu D---EHLLEKYLEGEE-LTVDEIKGAIRKLTIAEIIYVPLGSAFNKNGVQMLDAVVDY  
18nm S---EELMDKYLEGEE-LTEEEIVGALRQRTLAGEIQPMLGSAFNKNGVQMLDAVVEL  
19pae N---EELMNKYLEGEE-LSEAEIKEGLRLRLTACEIVPACGSSFAFNKNGVPLVLDVAVDY  
21rp -----  
22syn D---DALMEKYLEGEE-LTADELVAGLRRGTIAGTMVPLGSAFNKNGVQMLDAVVDY  
23tm D---EELMEKYLEGEE-ISEEIKRVLKATIENTKAVPVLGSAFNKNGVQMLDAVIDY  
24tp S---DELAEAYLEGTE-T-DQLIRAAVRAGTIAEKFPVPFCGSAFNKNGVQPLDAITYY  
25vib S---EELMEKYLEGEE-LSEVEIKQALRQRTINNEIVLAACGSAFNKNGVQAVLDVAVIEF  
26xyf T---EELMDKYLEGEE-LKEQEILQGLRERTLKVIEIPVFCGSAFNKNGVQMLDVITHL  
-----  
01aae LPSPIDLPPVKGTPNPKTGEEVHRHPSD--DEPFCAAYAFKVMSPDYAGQLTYIRVFSGLTKA  
04bb LPSPHIDKNTALDLNN--NEKEIDLKIDNSLPTVALAFKLEDGQY--GQLTYVIRVYQGLTKK  
05bs LPAPTDVAIKGRTPDTNNEIERHSSD--EPPFSALAFKVMTPDPVVGKLTFFRVYSGTLD  
06cj LPAPDEVANIKG-EYEDGTEVSVKSTDDGE--FAGLAFKIMTDPFVGQLTFFRVYRGCLLES  
07cp LPSPDLRGNIRGINLKTQDEISLEPRR--DGPLAALAFKIMTDPYVGRITFIIRVYSGTLKK  
08ct LPSPDLRGNVIRGINLKTGEEVSLKPSK--DGPLAALAFKIMTDPYVGRITFIIRVYSGTLKK  
09der LPSPLEVPISIKGTED--GETVTEFPADPEGLAALAFKIMADPYVGRITFIIRVYSGTLTS  
10ec LPSPVDVPAINGILDDGKDTAERHASDDEPFSAALAFKIAITDPFVGSLTFFRVYSGVNS  
11hi LPAPTDIPAIGKINPD--ETEGERHASDEEPFSSALAFKIAITDPFVGSLTFFRVYSGVNS  
12hp LPAPTEVVDIKGIDPKTEEEVFKSSDDGE--FAGLAFKIMTDPFVGQLTFFRVYRGKLES  
13mg LPSPVDVPPAKAIDAN--NKEISIKASD--DANFGLAFKVAITDPFVGRLTFIRVYAGVLKS  
15mp LPSPVDVPPAKAYGED--GNEVLISASD--DAPFVGLAFKVAITDPFVGRLTFIRVYSGVLKS  
17mtu LPSPDLVPPAIGHAPAKEDEEVVRKATTDPEFAALAFKIAITHPFFGKLTFIIRVYSGTVES  
18nm LPAPTDIPPVQGVNPN--TEEADSRQASDEEKFSALAFKMLNDKYVQGLTFIRVYSGVKS  
19pae LPAPTEIPAIGVSPDDETVEDERHADDNEPFSSALAFKIAITDPFVGSLTFFRVYSGVLS  
21rp -----  
22syn LPSPLEVPATIEGHLPD--GEVATRAPED--KAPLSALAFKVMADPF--GRLTFVRVYSGVLEK  
23tm LPSPDLPPVKGWRVSDGEVYRKPD--NEPFTALVFKVQVDPYIGKLVYFRVYSGRLEK  
24tp LPDPTVNTALDLDR--AEEPVTLSTADAPVVALGFKLEDGKY--GQLTYVIRVYQGTIKK  
25vib LPSPTDVPAIKGI--DDRENSVERHADDNEPFSSALAFKIAITDPFVGSLTFFRVYSGVNS  
26xyf LPSPADRPVQGLDEK--GNECRKASDSEPFSAALAFKIMTDPFVGSLTFFRVYSGVLS  
-----  
01aae GSYVYNATKDEKQKAGRLLLMHANSREEIQQVSAGEICAVVGLDA--ATGDTLCDEKHP-I  
04bb GQELINSRTSKKFKVGRILRMHANNNTIEFGSGDIVALFGIEC--ASGDTFCDSIN-Y  
05bs GSYVKNSTKGRKERVGRILQMHANSREEISTVYAGDIAAAVGLKDTTGTDLCDKDL-V  
06cj GSYAYNSTKDKKERISRLLEMHANSKREEIKVLVYAGEI GAVVGLKDTTGTDLASEDK-V  
07cp GSAILNSTKDKKERISRLLEMHANERTDRDEFTVGDIGACVGLKFSVTGDTLCDNQE-I  
08ct GSAILNSTKDKKERISRLLEMHANERTDRDEFTVGDIGACVGLKFSVTGDTLCDNQE-I  
09der GSYVYNASKDKRERVGRLLKMHANSREEVTELKAGELGAVIGLKADAGTGNLTIGDGDRLV  
10ec GDTVLNSVKAARERFGRIVQMANKREEIKEVRAGDIAAAIGLKVTTGDTLCPDPAP-I  
11hi GDTVLNSVRQKRERFGRIVQMANKREEIKEVRAGDIAAAIGLKVTTGDTLCAIDAP-I  
12hp GSYVYNSTKDKKERVGRLLKMHANSKREDIKEVYAGEICAFVGLKDTTGTDLCDKNA-V  
13mg GSYVKNVRKNKKERVSRVLKMHANRNEIDEIRAGDICAIVGLKDTTGTETLDDKID-V  
15mp GSYVKNVRKNKKERVSRVLKMHANRNEIDEIRAGDICAIVGLKDTTGTETLDDKID-V  
17mtu GSQVINATGKKERLGLKFMHNSKNPVDRAAGHIYAVIGLKDTTGTDLSDPNQ-I  
18nm GDTVLNSVKGTRERIGRLVQMTAADRTIEEVRAGDIAAAIGLKVTTGTETLCAESAP-I  
19pae GDSVLNSVKGKKERVGRMVQMHANQREEIKEVRAGDIAAALIGMKDVTGDTLCSIEKP-I  
21rp -----  
22syn GSYVLNSTKEKKERISRLIILKADDRIEVDQLNAGDLGAVLGLKDTTGTDLCDQEP-I  
23tm GSYVYNSTKGRERISRIVFMHADRKEEVYFRPGDIAAGVGLKVSQGTDLWHEKEP-I  
24tp GAELFNVRARKKFKVGRVLRMNSNQMEDISEGTPGDIVALFGVDC--ASGDTFCSGDLN-Y  
25vib GDAVYNSVKQKKERVGRIVQMANKRDEIKEIRAGDIAAAIGLKVTTGTETLCPDNHV-Y  
26xyf GDQVYNSVKLKKERVGRILQMHNSQRDEIKEVRAGDIAAAVGLKDVTTGDTLCPDNHI-I

```

01aae ILEKLEFPDPVISMAIEPKTKDQKLSQVLNFKMKEDPTFRATTPDTPETGQILIHGMGEL
04bb SMTSMFIPDPVISLSVKPKDKKSADNMALGRFTKEDPTFKTYVDIESNETI IQGMGEL
05bs ILESMEFPPEPVIDVAIEPKSKADQDKMGIALAKLAEDPTFRQTNPETGQTI ISGMGEL
06cj ILERMDFPDPVISVAVEPKTKADQEKMSIALNLKLAQEDPSFRVSTDEESGQTI ISGMGEL
07cp VLERIEFPDPVIDMAIEPKSKGDREKLAQALSSLEEDPTFRVSTNEETGQTI ISGMGEL
08ct VLERIEAPEPVIDMAIEPKSKGDREKLAQALSLEEDPTFRVSTNEETGQTI ISGMGEL
09der LLESIDIEPEPVIKLAIEPKTKADQEKMGVLQKLAEDPTFKVETDQESGQTTIAGMGEL
10ec ILERMEFPPEPVISIAVEPKTKADQEKMGALGRLAQEDPSFRVWTDDEESQTTIAGMGEL
11hi ILERMEFPPEPVISVAVEPKTKADQEKMGALGRLAQEDPSFRVHTDEESQTTI ISGMGEL
12hp VLERMEFPPEPVIIHIAVEPKTKADQEKMGVALGLAEEDPSFRVMTQETGQTLIGMGEL
13mg QLEAMQFAEPVISLAVEPKTKADQEKMSIALSKLAEDPTFKTFSDPETGQTTIAGMGEL
15mp QLEAMQFAQPVISLAVEPKTKADQEKMSIALSKLAEDPTFKTFDTPETGQTTIAGMGEL
17mtu VLESMTFPDPVIEVAIEPKTKSDQEKLSLSIQKLAEDPTFKVHLDSETGQTVIGMGEL
18nm ILERMEFPPEPVIIHIAVEPKTKADQEKMGIALNLAKEDPSFRVHTDEESQTTI ISGMGEL
19pae ILERMDFPPEPVISVAVEPKTKADQEKMGIALGLAQEDPSFRVKTDEESGQTTI ISGMGEL
21rp -----
22syn ILESLFVPQPVISVAVEPKTKQMDKLSKALQSLSEEDPTFRVSDPETNQTVIAGMGEL
23tm ILEKIDFPPEPVISLAVEPTKADDEKLVKALLALEEDPTLQVRVDKETGENI ISGMGEL
24tp AMTSMFVPEPVISLITPKDKRSADQVSKALNRFTKEDPTFRSFVDPESNQTTI IQGMGEL
25vib ILERMEFPPEPVIIHIAVEPRSKADQEKMGIALGLAEEDPSFRVETDAETGQTLISGMGEL
26xyf TLERMIFPEPVISMAVEPKTKSDQEKMGALGRLAQEDPSFRVKTDEESGQTTI ISGMGEL

01aae HLEIMVDRMKREYGIENVVGKPPQVAYKETIR--KKAIEGGEKFKIKTGGRGQYGHAIIEIEP
04bb HLEVYIERMKREFKAEEVETGMPQVAYRETIT--GKAEFNYTHKKQSGGAGQFGAVAGFMEP
05bs HLDIIVDRMKREFKVEANVGAPQVAYRETIR--TGAKVEGKFVRQSGGGRQYGHVVFIEFEP
06cj HLEIIVDRMLREFKVEAEVGGPPQVAYRETIR--KTVEQEYKYAKQSGGGRQYGHVFLRLEP
07cp HLDILDRMIREFKVEANVGKPPQVSYKETIT--VSGNSETKYVKQSGGGRQYAHVCLEIEP
08ct HLDILDRMIREFRVEANVGKPPQVSYKETIT--KTSNSETKYVKQSGGGRQYAHVCLEIEP
09der HLEILVDRLKREYKVEANVGAPQVAYRETIT--KQVEVDSKFARQSGGGRQYGHVKLRPEP
10ec HLDIIVDRMKREFKVEANVGKPPQVAYRETIRKQVTDVEGHAKQSGGGRQYGHVVIDMYP
11hi HLDIIVDRMKREFKVEANIGKPPQVSYRETITRNVNDVEGHAKQSGGGRQYGHVVIDLYP
12hp HLEIIVDRMLREFKVEAEIGQPPQVAFRETIR--SSVSKEHKYAKQSGGGRQYGHVFIKLEP
13mg HLDILVDRMKREFKVEANIGAPQVSFRETFR--STSEVEGKYIKQSGGGRQYGHVKIRFEP
15mp HLDILVDRMRREFKVEANVGAPQVSFRETFR--KESEVEGKYIKQSGGGRQYGHVKIRFEP
17mtu HLDILVDRMRREFKVEANVGKPPQVAYKETIKRLVQNVEYTHKKQSGGSGQFAKVINLEP
18nm HLEIIVDRMKREFGVEANIGAPQVAYRETIR--KAVKAIEYKHAKQSGGGRQYGHVVIEMEP
19pae HLDIIVDRMKREFGVEANIGKPPQVAYRETITKDNVEIEGKFVRQSGGGRQYGHVWIRFSA
21rp -----
22syn HLEILVDRMLREFKVEANVGAPQVAYRETIR--KAVQAEKGKFIQSGGGRQYGHVVIIEVEP
23tm HLEIIVDRMLREFGVNVRVGPPQVAYRETIR--KSAEAEKGKFIQSGGGRQYGHVILRIEIP
24tp HLDVYIERMRREYKCEVETGMPQVAYREAI--ARADFNYTHKKQSGGSGQFGRVAGFIEP
25vib HLDIIVDRMKREFGVDCNVGKPPQVAYRETIRGK--SEVEGKFVRQSGGGRQYGHVWLKIEP
26xyf HLDIIVDRMRREFKVEANVGKPPQVAYRETIRKSDVKSDYKHVKQSGGGRQYGHVVIIEISP

01aae LPRGA-----GFEFIDDIHGGVIPKEFIPSVSEKGVKEAMQNGILAGYPVVDVVRVLF
04bb LNKE-----GETYEFVNLIGGVIPTEYIPSCDKGFKQAMEKGLIGFPIVDIKITIN
05bs NEEG-----AGFEFENAI VGGVVPREYIPAVQAGLEDAENGLAGFLPIDIKAKLF
06cj LEP--GS-----GYEFVNDIKGGVIPKEYIPAVDKGVQALQNGVLAGYPVEDVKVTYV
07cp NEPGK-----GNEVVSIVGGVIPKEYIPAVIKGVEEGLNTGVLAGYGLVDVKVSIV
08ct NEPGK-----GNEVVSIVGGVIPKEYIPAVIKGVEEGLNSGVLAGYGLVDVKVSIV
09der LEPE-----AGFIFENAVVGGTVPKEYIGPAQKGVVEAMQSGPMLGFPVVDIKVVIY
10ec LEPSGNPK-----GYEFINDIKGGVIPGEYIPAVDKGIEQLKAGPLAGYPVVDGIRLH
11hi LDP--EGP-----GYEFVNEIKGGVIPGEYIPAVDKGIEQLKSGPLAGYPVVDLGVRLH
12hp KEP--GS-----GYEFVNEISGGVIPKEYIPAVDKGIEQAMQNGVLAGYPVVDKVTLY
13mg NKD-----KGFEFVDKIVGGRIPREYIKPVQGTGLENAMNSGPLAGYPMIDIKATLF
15mp NKD-----KGFEFVDKIVGGRIPREYIKPVQGTGLENAMASGPLAGYPMIDIKATLF
17mtu FTGEEGA-----TYEFESKVTGGRIPREYIPSVDAQAGDAMQGVGLAGYPLVNLKVTL
18nm MEP--GGE-----GYEFIDEIKGGVIPREFIPSVDKGIRDTLNGIAGYPVVDVIRLIV
19pae ADVDEKGN--ITEGLVFENEVVGVPKEYIPAIQKIEEQMKNGVAGYPLIGLKATVF
21rp -----
22syn TEPG-----TGFEFVSKIVGGVIPKEYIPASEQGMKEACASGVLAGYPVIDLKATLV
23tm IPEEE-----G---K---NF-----EFMPAIEAGIKEAMMAGPLAGYPVVRVRAIVL
24tp IA-----GQDYEFVDQIKGGVIPNEFIPSCDKGFRITAVKGTGLIGFPIVGVRTIN
25vib AEPGQ-----GVFVVDIAAGGVIPKEFINPVAKGIEEQMNGVLAGYPVLDVKATLF
26xyf MSVDVKQHPDVKGDFLFINETIGGVIPKEFISP IEKGLRETITTSGLAGFPVVGKVKLV

01aae DGSYHEVDSSDIAFQVAGSLAFKDAAKKADPVILLEPIMEVEVETPEKYVGDVIGDLNSRR
04bb DGQYHIVDSSDIAFQLAAIGAFREAYEKAAPTILEPIMKVTLEGTFFQGNMFGLLNQRR
05bs DGSYHVDVDSNEMAFKVAASMAKNAVSKCNPVILEPIMKVEVVIPEEYMGDIMGDITSRR
06cj DGSYHEVDSSSEMAFKLAASMGFKEGARKAGAVILEPMMKVEVETPEDYMGDVIGDLNRR
07cp FGSYHEVDSSSEMAFKICGSMVAVKDA CRKAKPVILEPIMKVAVITPEDHLDGVIGDLNRR
08ct FGSYHEVDSSSEMAFKICGSMVAVKDA CRKALPVILEPIMKVITVITPEDHLDGVIGDLNRR
09der DGSYHEVDSSSEMAFKIAGSMGLKAVQKGPALIEPVMRVEVTTPEEYMGDVIGDLNSRR
10ec FGSYHVDVDSSELAFLAASIAFKEGFKAKPVILLEPIMKVEVETPEENTGDVIGDLNSRR
11hi FGSYHVDVDSSELAFLAASLAFAKAAFSKANPVILLEPIMKVEVETPEVYVGDVIGDLNSRR
12hp DGSYHVDVDSSEMAFKIAGSMFKEASRAANPVILLEPMMKVEVEPEEYMGDVIGDLNRR
13mg DGSFHEVDSSSEMAFKIAASLALKEAGKQCNPVILLEPIMAEVTVPEQYFGDTMGDISSRR
15mp DGSFHEVDSSSEMAFKIAASLALKEAGKQVCPVILEPIMAEVTVPEQYFGDTMGDISSRR
17mtu DGAYHEVDSSSEMAFKIAGSQVLKAAALAPVILEPIMAVEVTTPEYMGDVIGDLNSRR
18nm FGSYHVDVDSQLAFELAASQAFKEGMRQSPALLEPIMAVEVETPEEYMGDVMDLNSRR
19pae DGSYHVDVDSNEMAFKIAASMATKQLAQKGGKVLEPIMKVEVVTPEYMGDVMDLNSRR
21rp -----
22syn DGSFHDVDSSEMAFKIAGSMATREAVQADPVILLEPVMKVEIPEVDFDMGNVIGDLNRR
23tm DGSYHEVDSSSEMAFKIAASMAFKEAMKAPVILEPIMKLEITTPPEEYMGNIISDLNSRR
24tp DGQSHPDVSSDIAFQAAAI GAFREAYNGAKPVILEPIMRVSVGEQFQEGSVFGLNQRR
25vib DGSFHDVDSSEMAFKIAGSMFKAAGALEAQPVILLEPIMKVEITTPEDVMDGVVDLNSRR
26xyf FGSYHVDVDSSEMAFKLAASMAFKQGAFAKANPVILLEPIMKVEIVSPEDVLDIMGDVSRRR

```



|       |                                                                |
|-------|----------------------------------------------------------------|
| 01aae | GKIMGMENK-GVITVIKAHVPLAEMFGYATTLRSLTQGRGTFIMKFSHYDEVPPQIAEKI   |
| 04bb  | GIITGSLED-GFSKVEAEVPLSEMFSTVLRSSQTQGAEFSMFLKYGKVPKSAIFDEL      |
| 05bs  | GRVEGMEAR-GNAQVVRAMVPLAEMFGYATALRSNTQGRGTFTHMDHYGEVPSVAEEI     |
| 06cj  | GQVNSMDEE-GGNKIITAFCLAEEMFGYSTDLRSQTQGRATYSMEFDHYDEVKNVADEI    |
| 07cp  | GKILGQESS-RGMAQVNAEVPLSEMFYTTSLRSLTSGRATSTMEPAFFAKVPQKIQEEI    |
| 08ct  | GKILGQESS-RNMAQVSAEVPLSEMFYMTSLRSLTSGRATSTMEPAFFAKVPQKIQEEI    |
| 09der | GQIQGMEAR-GNAQIVKAFVPLSEMFYATDMRSKTQGRASYSMFFDHYTQVPTNLAQQL    |
| 10ec  | GMLKGQESE-VTGKIIHAEVPLSEMFYATQLRSLTKGRASYTMEFLKYDEAPSNVAQAV    |
| 11hi  | AMVNGQEAN-EFVVKIYAEVPLSEMFYATDLRSQTQGRASYSMFLKYAEAPTSVAAAV     |
| 12hp  | GQINSMDDE-LGLKIVNAFVPLSEMFYSTDLRSATQGRGTYSMFHDHYGEVPSNIAKEI    |
| 13mg  | GIIEGTEQR-DNVQLIKAKVPLKEMFGYATDLRSFSQGRGNVVMQFSHYAETPKSVVNEI   |
| 15mp  | GLIEGTEQR-DNVQVIKAKVPLKEMFGYATDLRSFSQGRGNVVMQFSHYAETPKSVVNEI   |
| 17mtu | GQIQAMEER-AGARVVRHAEVPLSEMFYVGDLSKTQGRANYSVMFDSYSEVPANVSKEI    |
| 18nm  | GVVLGMDDDIGGKKVRAEVPLAEMFGYSTDLRSATQGRATYSMEFKKYSEAPAHIAAAV    |
| 19pae | GLIQGMEDTVSG-KVIRAEVPLGEMFGYATDVRSMQGRASYSMFSKYAEAPSNIVEAL     |
| 21rp  | -----                                                          |
| 22syn | GHIIEGQETE-QGIAKVAASVPLAEMFGYATDIRSKTQGRGIFSMEFSHYAEVPRNVAEAI  |
| 23tm  | AKVESLETR-CHLKIVIAKVPLSETFGYATVLRSLSQGRASYIMQFSHYQEVPEKIAEKI   |
| 24tp  | GVVSSADD-EQFSRVDAEVPLSEMFSTVLRSSQTQGAEYSMEFAKYKAGQCQVDTSL      |
| 25vib | GIIEGMDGEPAGLKI IHAKVPLSEMFYATDLRSATQGRASYMEFAEYADVPKNIADAI    |
| 26xyf | GVLLGQDDDS-LSGKVINAMIPLEGEMFGYATSLRSMQTQGRATFAMEFDHYEAPANTADTV |
| 01aae | IGERMAGKSS-----                                                |
| 04bb  | RKKFNRQNK-----                                                 |
| 05bs  | IKKNKGE-----                                                   |
| 06cj  | IKKRN-----                                                     |
| 07cp  | VKK-----                                                       |
| 08ct  | VKK-----                                                       |
| 09der | MKK-----                                                       |
| 10ec  | IEARGK-----                                                    |
| 11hi  | IEARKK-----                                                    |
| 12hp  | VEKRKG-----                                                    |
| 13mg  | IANKK-----                                                     |
| 15mp  | IATKK-----                                                     |
| 17mtu | IAKATGE-----                                                   |
| 18nm  | TEARKG-----                                                    |
| 19pae | VKKQG-----                                                     |
| 21rp  | -----                                                          |
| 22syn | VAKSRGYA-----                                                  |
| 23tm  | IKVV-----                                                      |
| 24tp  | IKKEYQEKRAEQRLTACVCLITIGRKIPDLRARGGTQEDAKPVIKCWFQGVQGPLSSVR    |
| 25vib | IAEHG-----                                                     |
| 26xyf | IKKT-----                                                      |
| 01aae | -----MAKEKFE                                                   |
| 04bb  | -----MKFRRLVMAKEVFQ                                            |
| 05bs  | -----MAKEKFD                                                   |
| 06cj  | -----MAKEKFS                                                   |
| 07cp  | -----MSKETTFQ                                                  |
| 08ct  | -----MSKETTFQ                                                  |
| 09der | -----MAKGTFE                                                   |
| 10ec  | -----MSKEKFE                                                   |
| 11hi  | -----MSKEKFE                                                   |
| 12hp  | -----MAKEKFN                                                   |
| 13mg  | -----MAREKFD                                                   |
| 15mp  | -----MAREKFD                                                   |
| 17mtu | -----VAKAKFQ                                                   |
| 18nm  | -----MAKEKFE                                                   |
| 19pae | -----MAKEKFE                                                   |
| 21rp  | -----MAKAKFE                                                   |
| 22syn | -----MARAKFE                                                   |
| 23tm  | -----MAKEKFE                                                   |
| 24tp  | NALGSRAGCARCLRGPFPGCSAVDPSFRFFVPPCCVRACGVRRICATDLLRRIFMAKEKFA  |
| 25vib | -----MSKEKFE                                                   |
| 26xyf | -----MAQDKFK                                                   |
| 01aae | RTKEHVNVTIGHVDHGKSTLTAITCVLAAGLVE-GGKAKCFKYEIDKAPEEKERGIT      |
| 04bb  | RTKPHMNVGTIGHVDHGKSTLTAATISYCSLKNKDA----KALKYEDIDNAPEEKERGIT   |
| 05bs  | RSKSHANIGTIGHVDHGKSTLTAATITVLHKS----GKGTAMAYDQIDKAPEERERGIT    |
| 06cj  | RNKPVNIGTIGHVDHGKSTLTAATISAVLSRR----GLAELKDYDNIDNAPEEKERGIT    |
| 07cp  | RNKPVINIGTIGHVDHGKSTLTAATITRSLG----DGLASFRDYSSIDNTPEEKERGIT    |
| 08ct  | RNKPVINIGTIGHVDHGKSTLTAATITRSLG----DGLADFRDYSSIDNTPEEKERGIT    |
| 09der | RTKPHVNIGTIGHVDHGKSTLTAATITFAASAD----PTIETLAYDQIDKAPEEKERGIT   |
| 10ec  | RTKPHVNVTIGHVDHGKSTLTAATITVLAKT----YGGAARAFDQIDNAPEEKERGIT     |
| 11hi  | RTKPHVNVTIGHVDHGKSTLTAATITVLAKH----YGGAARAFDQIDNAPEEKERGIT     |
| 12hp  | RTKPHVNIGTIGHVDHGKSTLTAATISAVLSK----GLAEMKDYDNIDNAPEEKERGIT    |
| 13mg  | RSKPHVNVTIGHVDHGKSTLTAATITVLAKT----GKSAATRYDIDKAPEEKERGIT      |
| 15mp  | RSKPQLNVGTIGHVDHGKSTLTAATITVLAKT----GKSAATRYDIDKAPEEKERGIT     |
| 17mtu | RTKPHVNIGTIGHVDHGKSTLTAATITVLAKT----DLNETKAFDQIDNAPEEKERGIT    |
| 18nm  | RSKPHVNVTIGHVDHGKSTLTAATITVLAKT----FGGAAKAYDQIDNAPEEKERGIT     |
| 19pae | RNKPVNVTIGHVDHGKSTLTAATITVLAKT----WGGSARAFDQIDNAPEEKERGIT      |
| 21rp  | RTKPHVNIGTIGHVDHGKSTLTAATITVLAKT----GGAKATAYDQIDNAPEEKERGIT    |
| 22syn | RTKPHVNIGTIGHVDHGKSTLTAATITVLAKT----GGAKATAYDQIDNAPEEKERGIT    |
| 23tm  | RTKPHVNVTIGHVDHGKSTLTAATITVLAKT----KGLAQYIPYDQIDKAPEEKERGIT    |
| 24tp  | RTKPHVNVTIGHVDHGKSTLTAATITVLAKT----KGLAQYIPYDQIDKAPEEKERGIT    |
| 25vib | RTKPHVNVTIGHVDHGKSTLTAATITVLAKT----YGGKARDFASIDNAPEEKERGIT     |
| 26xyf | RTKLHVNVTIGHVDHGKSTLTAATITVLAKT----FGGEKAYDAIDNAPEEKERGIT      |

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01aae | INITHVEYETAKRHYAHVDCPGHADYIKNMITGAAQMDGAILVVSADGPMPTREHVLL      |
| 04bb  | INARHIEYETANRHYAHVDCPGHADYIKNMITGAAQMDAAILVVAADSGAEPQTKHELLL    |
| 05bs  | ISTAHVEYETETRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL      |
| 06cj  | IATSHIEYETDNRRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL     |
| 07cp  | INASHVEYETPNRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSATDGAMPQTKHEILL    |
| 08ct  | INASHVEYETANRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSATDGAMPQTKHEILL    |
| 09der | INTAHVEYQTETRHYSHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL      |
| 10ec  | INTSHVEYDTPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVVAATDGPMPQTKREHILL   |
| 11hi  | INTSHVEYDTPTRHYAHVDCPGHADYVKNMITGAAQMDGAILVVAATDGPMPQTKREHILL   |
| 12hp  | IATSHIEYETENRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL      |
| 13mg  | INSAHVEYSSDKRHYAHVDCPGHADYIKNMITGAAQMDGAILVVSATDSVMPQTKREHILL   |
| 15mp  | INTAHVEYSSDKRHYAHVDCPGHADYIKNMITGAAQMDGAILVVSATDSVMPQTKREHILV   |
| 17mtu | INTAHVEYQTDKRHYAHVDCPGHADYIKNMITGAAQMDGAILVVAATDGPMPQTKREHVLL   |
| 18nm  | INTSHVEYETETRHYAHVDCPGHADYVKNMITGAAQMDGAILVCSAADGPMPTREHILL     |
| 19pae | INTSHVEYDSAVRHYAHVDCPGHADYVKNMITGAAQMDGAILVCSAADGPMPTREHILL     |
| 21rp  | ISTAHVEYETQNRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL      |
| 22syn | INTAHVEYETDSRHYAHVDCPGHADYVKNMITGAAQMDGAILVVSADGPMPTREHILL      |
| 23tm  | INITHVEYETEKRHYAHIDCPGHADYIKNMITGAAQMDGAILVVAATDGPMPQTKREHVLL   |
| 24tp  | INTRHLEYQSDRRHYAHIDCPGHADYVKNMITGAAQMDGAILVVSAPDGVMPQTKHEILL    |
| 25vib | INTSHVEYDTPNRHYAHVDCPGHADYVKNMITGAAQMDGAILVVAATDGPMPQTKREHILL   |
| 26xyf | ISTAHVEYETEVRRHYAHVDCPGHADYVKNMITGAAQMDGAILVCSAADGPMPTREHILL    |
|       |                                                                 |
| 01aae | ARQVNVPPYIVVFMNKCMDVD--DEELLELEVEVRELLSKYEYFPGDEVPIRGSALGALQ    |
| 04bb  | AQRMGIKKIIVFLNKLDLAD--PELVLEVEVEVLEVEKYGFS--ADTPIKGSAFGAMS      |
| 05bs  | SKNVGVPPYIVVFLNKCMDVD--DEELLELEVEMVRDLLSEYDFPGDDVPVVKGSALKALE   |
| 06cj  | SRQVGVPPYIVVFMNKCMDVD--DAELLELEVEMIRELLSSYDFPGDDTPIISGSALKALE   |
| 07cp  | ARQVGVPPYIVVFLNKCMDVD--PELVLEVEVEVLEVEKYGFS--CPIIRGSALKALE      |
| 08ct  | ARQVGVPPYIVVFLNKCMDVD--PELVLEVEVEVLEVEKYGFS--CPIIRGSALKALE      |
| 09der | ARQVGVPPYIVVFMNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDLPVVKGSALRALE  |
| 10ec  | GRQVGVPPYIVVFLNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDTPIIRGSALKALE  |
| 11hi  | GRQVGVPPYIVVFLNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDTPIIRGSALQALN  |
| 12hp  | SRQVGVPHIIVVFLNKCMDVD--DQELLELEVEMIRELLSKYEYFPGDDTPIVAGSALKALE  |
| 13mg  | ARQVGVPKMVVFLNKCMDIAS--DEEVQELVAEEVRDLLTSYGFDDGKNTPIIYGSALKALE  |
| 15mp  | ARQVGVPRMVVFLFKCDIAT--DEEVQELVAEEVRDLLTSYGFDDGKNTPIIYGSALKALE   |
| 17mtu | ARQVGVPPYILVALNKADAVD--DEELLELEVEMIRELLSKYEYFPGDDLPVVKGSALRALE  |
| 18nm  | ARQVGVPPYIVVFMNKCMDVD--DAELLELEVEMIRELLSKYEYFPGDDCPVVKGSALKALE  |
| 19pae | SRQVGVPPYIVVFLNKCMDVD--DAELLELEVEMIRELLSKYEYFPGDDTPIIISGSALMALE |
| 21rp  | AKQVGVPMVFLNKCMDVD--DPDLELEVEMIRELLSKYEYFPGDDTPIIISGSALKALE     |
| 22syn | AKQVGVPKLVVFLNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDTPIIISGSALKALE  |
| 23tm  | ARQVGVPPYIVVFMNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDTPIIISGSALKALE |
| 24tp  | ARQVGVPSIIVFLNKCMDVD--PELLELEVEMIRELLSKYEYFPGDDTPIIISGSALKALE   |
| 25vib | GRQVGVPPYIVVFMNKCMDVD--DEELLELEVEMIRELLSKYEYFPGDDLPVVKGSALGALN  |
| 26xyf | ARQVGVPPYIVVFLNKCMDVD--DAELLELEVEMIRELLSKYEYFPGDDTPIIISGSALKALE |
|       |                                                                 |
| 01aae | ELEQ----NSP-GKWWESIKELLNAMDEYIPTPREVDKPFMLPIEDVFSISGRGTVVTG     |
| 04bb  | NP-----EDPESTCKVKELLESMDNYFDLPERDIDKPFLLAVEDVFSISGRGTVATG       |
| 05bs  | G-----DAEWAKIFELMDAVIDEYIPTPERDTEKPFMPVEDVFSITGRGTVATG          |
| 06cj  | EAK-----AGQDGEWSAKIMDLMAVDSYIPTPTDTEKDFLMPIEDVFSISGRGTVVTG      |
| 07cp  | GD-----ANYIEKVRELMAQVDDNIPTPERIDKPFMLPIEDVFSISGRGTVVTG          |
| 08ct  | GD-----AAYIEKVRELMAQVDDNIPTPERIDKPFMLPIEDVFSISGRGTVVTG          |
| 09der | ALQSNPKMARGTDKWWDYIWELLDVDSYIPTPERDTEKDFLMPVEDVFTITGRGTVATG     |
| 10ec  | G-----DAEWAKIFELAGFLDSYIPEPERAIDKPFLLPIEDVFSISGRGTVVTG          |
| 11hi  | G-----VAEWEEKILELANHLDYIPEPERAIDKPFLLPIEDVFSISGRGTVVTG          |
| 12hp  | EAK-----AGNVGEWGEKVLKMAEVDYIPTPERDTEKDFLMPVEDVFSIAGRGTVTG       |
| 13mg  | G-----DPKWEAKIHDLIKAVDEWIPTPTREVDPKPFLLAIEDTMTITGRGTVVTG        |
| 15mp  | G-----DPKWEAKIHDLMNAVDGQFLNVKWNPFLLAIEDTMTITGRGTVVTG            |
| 17mtu | G-----DAKWWASVEELMNAVDSEIPDPVRETDPKPFMLPVEDVFTITGRGTVVTG        |
| 18nm  | G-----DAAYEEKIFELAAALDSYIPTPERAIDKPFLLPIEDVFSISGRGTVVTG         |
| 19pae | GK-----DDNGIGVSAVKQLVETLDSYIPEPRAIDKPFLLPIEDVFSISGRGTVVTG       |
| 21rp  | G-----KPEGEKAINELMNAVDYIIPQIELQDKPFMLPIEDVFSISGRGTVVTG          |
| 22syn | G-----EKEYKDAILELMNAVDYIDTPEREVDKPFMLMAVEDVFSITGRGTVATG         |
| 23tm  | AP-----NDPNHEAYKPIQELLDAMDNYIPDPQRDVDPKPFMLPIEDVFSITGRGTVVTG    |
| 24tp  | DG-----ASPEDAACIEELLAAMDSYFEDPVRDDARPFLLSIEDVYITISGRGTVVTG      |
| 25vib | G-----EAQWEAKIVELAEALDYIPEPERAIDKPFMLPIEDVFSIQRGTVVTG           |
| 26xyf | G-----DQSEIGVPAIIRLAEALDTHIIPPERAIDKPFMLPVEDVFSISGRGTVVTG       |
|       |                                                                 |
| 01aae | RVERGVLRPGDEVEIVGLREEPLKTVATSIEMFRKVLDEALPGDNIGVLLRGVKGDDVER    |
| 04bb  | RIERGIVKVGDEVEIVGIGK-ETRKTTVTGVEMFRKILEQQAGDNVGLLLRGVDKKEIER    |
| 05bs  | RVERGVKVGDEVEIVGLQENKTTVTGVEMFRKLLDYAEAGDNIGALLRGVSRREEIQR      |
| 06cj  | RIEKGIVKVGDTIEIVGIGK-DTQTTTGTGVEMFRKEMDQGEAGDNVGLLLRGTKKEEIVR   |
| 07cp  | RIERGIVKVSQKVLVGLG-ETKETITVTGVEMFRKELPEGRAGENVGLLLRGIGKNDVER    |
| 08ct  | RIERGIVKVSQKVLVGLR-DTKETLLGLEMFKNQSKVRAGENVGLLLRGIGKNDVER       |
| 09der | RVERGTVKVGDEVEIVGLT-DTRKTTVTGIEMHFKLLDSGMAAGDNVGLLRGVARDVVER    |
| 10ec  | RVERGIKVGEEVEIVGIGK-ETQKSTCTGVEMFRKLLDEGRAGENVGLLLRGIKREEIER    |
| 11hi  | RVERGIIRTGDEVEIVGIGK-DTAKTTVTGVEMFRKLLDEGRAGENIGALLRGTKREEIER   |
| 12hp  | RIERGIVKVGDEVEIVGIGK-PTQKTTVTGVEMFRKELKEGEAGDNVGLLRGTKEEVER     |
| 13mg  | RVERGELKVGDEVEIVGLK-PIRKAVVTGIEMFKKELDSAMAGDNAGVLLRGVERKEVER    |
| 15mp  | RVERGELKVGDEVEIVGLR-PIRKAVVTGIEMFKKELDSAMAGDNAGVLLRGVDRKEVER    |
| 17mtu | RVERGVINNVEEVEIVGIRPSTTKTTVTGVEMFRKLLDQQAAGDNVGLLLRGVKREDVER    |
| 18nm  | RVERGIHVGDIEIVGLK-ETQKTTCTGVEMFRKLLDEGQAAGDNVGLLRGTKREDVER      |
| 19pae | RVERGIKVGEEVEIVGIGK-ATTKTTCTGVEMFRKLLDEGRAGENVGLLRGTKREDVER     |
| 21rp  | RVESGIKVGEEVEIVGLK-NQKTTCTGVEMFRKLLDEGQSGDNVGLLRGTKEEVER        |
| 22syn | RIERGIVKVGEEVEIVGIGK-DTRKATVTGVEMFRKLEEGMAAGDNVGLLLRGIKQEDIER   |
| 23tm  | RIERGIRIPGDEVEIVGLSIEYIKKTVTGVEMFRKELDEGIAAGDNVGLLRGTKEEVER     |
| 24tp  | RIECGVISLNNEVEIVGIGK-PTKTTVTGIEMFKLLDQQAAGDNVGLLRGVDRKEEVER     |
| 25vib | RIERGILKVGDEVAIVGIGK-ETVKTCTGVEMFRKLLDEGRAGENVGLLRGTKEEVER      |
| 26xyf | RVECGVIVKVGDEVEIVGIGK-PTSKTTVTGVEMFRKLLDQQAAGDNAGVLLRGTKREDVER  |

```

01aae GQVLAQPGSVKAHKRFRAQVYVLSKEEGGRHTPPFFVNYRPQFYFRITADVTGTGTVKLP---
04bb GQVLSAPGTITPHKKFKASIYCLTKEEGGRHKPPFGYRPQFFFRITDVTGVVAL-----
05bs GQVLAKPGTITPHSKFKAIEVYVLSKEEGGRHTPPFSNYRPQFYFRITDVTGIIHLP----
06cj GMVLAKPKSITPHTDFEAEVYVLSKDEEGGRHTPPFNRYRPQFYFRITDVTGSIKLA----
07cp GMVVCQPNSVKPHTKFKSAVYVLSKEEGGRHKPPFGYRPQFFFRITDVTG-VVTLP---
08ct GMVVCLPNSVKPHTRFKCAVYVLSKEEGGRHKPPFGYRPQFFFRITDVTG-VVTLP---
09der GQVLAKPGSIKPHTKFEASVYVLSKDEEGGRHSAFFGGYRPQFYFRITDVTGVVELQ----
10ec GQVLAKPGTIKPHTKFESEVYVLSKDEEGGRHTPPFKGYRPQFYFRITDVTGTIELP----
11hi GQVLAKPGSITPHTDFESEVYVLSKDEEGGRHTPPFKGYRPQFYFRITDVTGTIELP----
12hp GMVLCKPGSITPHKKFEGEIVYVLSKEEGGRHTPPFTNYRPQFYFRITDVTGSIKLA----
13mg GQVLAKPGSIKPHKKFKAIEYALKKEEGGRHTGFLNGYRPQFYFRITDVTGSIKLA----
15mp GQVLAKPGSIKPHKKFKAIEYALKKEEGGRHTGFLNGYRPQFYFRITDVTGSIKLA----
17mtu GQVVTKPGTITPHTFEFQVYVLSKDEEGGRHTPPFNRYRPQFYFRITDVTGTVVTLP----
18nm GQVLAKPGTITPHTKFAIEVYVLSKEEGGRHTPPFANYRPQFYFRITDVTGAVTLE----
19pae GQVLAKPGTIKPHTKFECEVYVLSKEEGGRHTPPFKGYRPQFYFRITDVTGNCEL---
21rp GQVLAKPGSIKPHDKFEAEVYVLSKEEGGRHTPPFTNYRPQFYFRITDVTGTIKLP---
22syn GQVLAKPGSITPHTFEFQVYVLSKDEEGGRHTPPFANYRPQFYFRITDVTGTIKSYTADD
23tm GQVLAAPGSIKPHKRFKAQIYVLSKDEEGGRHTPPFTGYRPQFYFRITADVTGEIVGLP---
24tp GQVLSKPGSIKPHTKFEAQIYVLSKDEEGGRHSPFFGYRPQFYFRITDITGTISLP----
25vib GQVLAKPGSITPHTKFESEVYVLSKDEEGGRHTPPFKGYRPQFYFRITDVTGSIKLA----
26xyf GQVLAKPGSIKAKHFEAEVYVLSKEEGGRHTPPFNGYRPQFYFRITDITGKVCPL----

01aae -EGVEMVMPGDNVELEVELIAPVALEEGLRFAIREGGRTVGAGVVTIKILD-----
04bb -EGKEMVMPGDNVDIIVELISSIAMDKNVEFAVREGGRTVASGRILEILE-----
05bs -EGVEMVMPGDNTEMNVELISTIAIEGTRFSIREGGRTVSGGVVSTITE-----
06cj -DGVEMVMPGENVRI TVSLIAPVALEEGLRFAIREGGRTVSGGVVSIKII-----
07cp -EGTEMVMPGDNVELDVELIGTVALEEGMRFAIREGGRTIGAGTISKINA-----
08ct -EGVEMVMPGDNVEFEVQLISPVALEEGMRFAIREGGRTIGAGTISKIIA-----
09der -EGVEMVMPGDNVTFVTELKPIAMEEGLRFAIREGGRTVGAGVVSIVLE-----
10ec -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVAIKLS-----
11hi -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVAIKII-----
12hp -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSNIIE-----
13mg -ENTEMVLPGDNASITVELIAPVACEGSKFSIREGGRTVGAGVVSNIIE-----
15mp -ENTEMVLPGDNTSISVELIAPVACEGSKFSIREGGRTVGAGVVSNIIE-----
17mtu -EGTEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVAIKII-----
18nm -EGVEMVMPGDNVITVTELKPIAMEEGLRFAIREGGRTVGAGVVSIVIA-----
19pae -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVAIKII-----
21rp -SDKQVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----
22syn GSAVEMVMPGDRIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----
23tm -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----
24tp -EGVDMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----
25vib -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----
26xyf -EGVEMVMPGDNVIMVVTLIHPIAMDDGLRFAIREGGRTVGAGVVSIVIA-----

01aae -----MSEQQVQLEKKWYALQVEPGKEN
04bb -----MSRAWYVVQTSYQYK
05bs -----MEKNWYVVHTYSGYEN
06cj -----MSTHKWYAIQTYAGSEM
07cp -----MYKWYVVQVFTAQEK
08ct -----MFKWYVVQVFTAQEK
09der -----MSIEWYAVHTYVGQED
10ec -----MSEAP-----KKRWYVVQAFSGFEG
11hi -----MTEETTVSKRWYVLAQFSGFES
12hp -----MMDWYAIQTYSGSEQ
13mg -----MQASELTPKWYVAPVSIKDEA
15mp -----MEQVELIPQWYVAPVSVKDEA
17mtu VTTFDGDTSAGEAVDLTEANAFQDAAPAEVDPAAALKAELRSKPGDWVVSYSAGYEN
18nm -----MSKKWYVVQAFSGFEK
19pae -----MAKRWYVVHAYSGYEK
21rp -----VTEQIDNILPASKNNVQWYVVHTASGAEK
22syn -----MSFTDDQSPVAEQNKTPSEGHWFVAVQVAGCEK
23tm -----MKKKWYIVLTMSCYEE
24tp -----MAKEWYILHTFSGREA
25vib -----MSEAP-----KKRWYVVQAFSGFEG
26xyf -----MKRWYVVHAYSGFEK

01aae EAKENLLKVLEGLK-DLVDEVIVP---A-----
04bb KIEQDIRLLINEGVFC-GVVLVVKAP-----
05bs KVKANLEKRVESMGQM-DKIFRVVVP-----
06cj AVKRAIENLVKDNIE-EQLKEIVVP-----
07cp KVKKALEDFKESGMT-DFIQEILP-----
08ct KVKKLEDFKESGMT-DFIQEILP-----
09der RVQDQLMDRRLGMYRTKIFQVLQP-----
10ec RVATSLREHIKLHNME-DLFGVEMVP-----
11hi RVALTLREYIKQQME-DQFGEVLVP-----
12hp SVKKAIENLANDHNIR-DRIQEIIVP-----
13mg VVK-NLKAKIQLGPN-HEIVDVVKLKEREV-----
15mp VVR-NLKAKVKGALFD-NEILDVRLKEREV-----
17mtu KVKANLETRVQNLVVG-DYIFQVEVP-----
18nm NVQRILEERIAREEMG-DYFGQILVP-----
19pae HVMRSIERVKLACME-EEFGEILVP-----
21rp RIKEDILRKIAKQMT-DFFEDILP-----
22syn RVKLNLEQRIHTLDA-DRIQEIIVP-----
23tm KVKENIEKKVEATGK-NLVGRIVPEEVLDATSPSERLILSPKAKLHVNGKDVNGGD
24tp RVERAVRMLVEHARIPTNVIFDIKIP-----
25vib RVAQSLREHIKMHGME-ELFGEVLVP-----
26xyf SVAQALDRDRISRIEQ-DRFGDVLVP-----

```

```

01aae -----
04bb -----
05bs -----
06cj -----
07cp -----
08ct -----
09der -----
10ec -----
11hi -----
12hp -----
13mg -----
15mp -----
17mtu -----
18nm -----
19pae -----
21rp -----
22syn -----
23tm -----
24tp -----
25vib -----
26xyf -----

01aae -----EEKVIRAQGEKYRLSLKGNARDISVLGKKGVITTFRIENGVEKVVES
04bb -----IEKVEEIRNGKKRIR-----
05bs -----EEEEEDIKNGKKKV-----
06cj -----TEDVIEF-KNGKEKI-----
07cp -----IENVMEVKKGEHKV-----
08ct -----SENVMEVKKGEHKI-----
09der -----EEEAVEIQEGGKKVN-----
10ec -----TEEVVEIRGGQRRK-----
11hi -----TEEVVENVAGRRK-----
12hp -----TEDIEVSKKSKTKV-----
13mg -----HEEVYSLKSGKLPRLKNT-----
15mp -----IEEVFSLKSGKLPRLKNT-----
17mtu -----TEEVTEIKNGQRRK-----
18nm -----VEKVVDIRNGRRTI-----
19pae -----TEEVVEMRNGQRRK-----
21rp -----VFGVSEVKRGKNVK-----
22syn -----KTPIVKIRKDGARVQ-----
23tm -----PLSKNEEYICELDGKIVEIERMKNVQTPDGEQDVVYIPLDVFDRIKKGKVKQGM
24tp -----EELLTEVKDGKRV-----
25vib -----TEEVVEMRAGQRRK-----
26xyf -----AEVEMRSGQRR-----

01aae -----VEGDTCVNAPPISKPGQKITCKENKTEAKIVLDNKIPPGYILIKAHMNDKL-----
04bb -----ERKIWPGYILIELDLPEVG-----WKDI
05bs -----VKKKVFPGYVLVEIVMTDD-----S
06cj -----SERSLYSGYVFALLDLNTEL-----
07cp -----VEKYIWPGYLLVKMHLTDES-----
08ct -----VEKYIWPGYLLVKMHLTDES-----
09der -----VKRKLPGYVVFQMDVEDDDA-PGELGES
10ec -----SERKFFPGYVLVQVMNDAS-----
11hi -----SERKFFPGYVLVEMMNDET-----
12hp -----TERSLYPGYVFIKVDLDTVL-----
13mg -----TFNKWFVLDDYRLRVKISEKNLLGRYIYIKMIYSEDA-----
15mp -----AFTKWFVLDEDRLVKISEKNLLGRYIYIKMIYSEDA-----
17mtu -----VNRKVLPGYILVRMDLTDD-----S
18nm -----SERKSYPGYVLVEMMTDSD-----
19pae -----SERKFFPGYVLVQVMEN-EG-----T
21rp -----VEKKLMPSYILIKMMTDKS-----
22syn -----GEEKIFPGYVLIRMIMDD-----A
23tm -----LAEARKFFAKVSGRVEVDYSTRKEIRIYKTKRRKLPFGYVFEVIMNDEA-----
24tp -----VRRKFFPGYLLVEMDLPEVDWR-----IV
25vib -----SERKFFPGYVLVQIMNDES-----
26xyf -----SEHKFFPGYVLQIETYEGGVPRIDNEC-----

01aae -----LMAIEKTPHVFRPVM-----VGGKPVPLKEEEVQNILNQIKR-----
04bb -----IANIKVQGVINFG-----VSKGQRPPIINDEEVKSVFMTGEIK-----
05bs -----WYVVRNTPGVTGFVGS-----AGSGKPTPLLPGEAETILKRMG-----
06cj -----WHRIQSLPKVGRFIG-----ESKKPTPLTEKDINLILEKVHN-----
07cp -----WLYVKSTAGIVEFLG-----GGVPVALSEDEVRSILTDIEE-----
08ct -----WSYVKKTQGVVEFLG-----GGAPVALSEEVKNILADLEE-----
09der -----WEVVRGTSGVTGFVG-----TATRVPVLSPEEQRLLTSGV-----
10ec -----WHLVRSVPRVMGFIG-----GTSRPPAPISDKEVDAIMNRLQ-----
11hi -----WHLVKSVPVMGFIG-----GTPDKPAPISKREADTILNRLEQ-----
12hp -----WHKIQSLPRVSRFIG-----ENKKPTPLSEADIGHILEKMN-----
13mg -----WRIVRNFPGITGIVGS-----SGRGALPIPLDEKADNNEQMLKGISINPSKRIMLT
15mp -----WRIIRNFPGITGIVGS-----SGRGALPTPLDQADNNEQMLKGISVNPKKRVLVT
17mtu -----WAAVRNTPGVTGFVG-----ATSRPSALALDDVVKFLLPRGSTRKAAGAATA
18nm -----WHLVKSTPRVSGFIG-----GRANRPTPISQREAEIILQVQT-----
19pae -----WHLVKDTPRVMGFIG-----GTADKPAPITDREADAILRRVAD-----
21rp -----WHLVKNIPGVTGFLG-----SKIVPKALTESEIQNIFNNLEA-----
22syn -----WQVVKNTPHVINFGVSEQRHYGRGRGHVLPMLSHGEVERIFRHVD
23tm -----YNFVRSPVPMGFVS-----SGGPVPVVKDREMRPILRLAGLEEYE-----
24tp -----CNEVRNIPGVSGFLGS-----SGNAKPQAVSADEARRILQKAGEIK-----
25vib -----WHLVRSVPRVMGFIG-----GTSRPPAPITDKEADAILNRLEK-----
26xyf -----WHLVKETPKVMGFIG-----GTADRPLPISSEADAILRRVQD-----

```

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01aae | -----GV                                                         |
| 04bb  | -----ANKS                                                       |
| 05bs  | -----MDE                                                        |
| 06cj  | -----RAA                                                        |
| 07cp  | -----KKSG                                                       |
| 08ct  | -----KKSG                                                       |
| 09der | -----VAAQPVVEEA                                                 |
| 10ec  | -----VGDK                                                       |
| 11hi  | -----NTDK                                                       |
| 12hp  | -----RAA                                                        |
| 13mg  | NTAIEMDSKDFEFQYILKQKQAIQK--PKEDESEIVDAEKLKEAFKLLQNSQEQDE        |
| 15mp  | NTAIVEMDADKFDEKCYILKHKQVKPEAIAQVNESGEIIDTNQFAQLMEA-NKAEQDE      |
| 17mtu | -----AAAAAGGLER                                                 |
| 18nm  | -----GIEK                                                       |
| 19pae | -----SGDK                                                       |
| 21rp  | -----EAKV                                                       |
| 22syn | -----EQE                                                        |
| 23tm  | -----EKKKP                                                      |
| 24tp  | -----CDRT                                                       |
| 25vib | -----ASEA                                                       |
| 26xyf | -----GAEK                                                       |
| 01aae | KPSKVEFEKGDQVRVIEGPFMNFVTGTVEEVHPEKRKLTVMISIFGRMTPV-ELDFDQVEK   |
| 04bb  | IFMLYDFEGERVRKGGPFDSFEGLISSIDYERKKLKVAQVIFGRSTPV-EVDFQHIEK      |
| 05bs  | RKTDIDFELKETVKVINDGPFANFTGSGIEEIDYKSKVVFVNMFGRETPV-ELEFTQIDK    |
| 06cj  | PKPKISFEEGENVRITEGPFANFTAIVEEYDMVRGLLKLNVISIFGRSTPV-EILYSQVEK   |
| 07cp  | VVQKHQFEVGSRVKINDGVFVNFVGTVEVFDKGRLSVMVISIFGRETRVDDLEFWQVEE     |
| 08ct  | VVQKHQFEVGSQVKINDGVFVNFVGTVEVFDKGRLSVMVISIFGRETRVDDLEFWQVEE     |
| 09der | PRIKVDLKAGDMVRVTSGPFADFGVGVSEVNPQAKVKVLVSIFGRSTPV-ELDFSQVAK     |
| 10ec  | PRPKTLFEPGEMVRVNDGPFADFGVGVSEVNDYKSRKLVSVISIFGRATPV-ELDFSQVEK   |
| 11hi  | PRHRNEYHPGEEVRVTEGPFADFGVGVSEVNDYKSRKLVSVISIFGRATPV-ELEFGQVEK   |
| 12hp  | PKPKIFFEQEVVRVVEGPFANFTATVEEYDVEHRKLKLNVISIFGRNTPI-EILHSQVEK    |
| 13mg  | WKEKATIIQSEQTKLDPSVLVPFLGKYEILDNDKVEQLFEFSVGNLVEVHLDTDTHVQG     |
| 15mp  | WNEDVAIVKSEANKVDPVSVLIPYLGKYEIVEGDTKVDLQKQFSVGNLVEVHLDGAIHQG    |
| 17mtu | PVVEVDYEVGESVTVMGPFATLPATISEVNAEQKQLKVLVSIFGRSTPV-ELTFGQVSK     |
| 18nm  | PKPKVEFEVGGQVRVNEGPFADFGVGVSEVNYERNKLRVSVQIFGRSTPV-ELEFSQVEK    |
| 19pae | PKPKTLFEPGETVRVINDGPFADFGVGVSEVNYEKSRIQVAVLIFGRSTPV-ELEFSQVEK   |
| 21rp  | AKNSKLYEVGEIVTVTDGPFETFMGTVEAIDKARNRLKVSVISIFGKATPI-ELNFNQVKK   |
| 22syn | PVVKIDMEIGDHIMVLGPFDFEGVDIEVSPERSKLLKALLSIFGRETPV-ELEFTQVEK     |
| 23tm  | VKVELGFKVGMVKIISGPFEDFAGVIKEIDPERQELKVNVTIFGRETPV-VLHVSEVEK     |
| 24tp  | PRIAQTFLVGQVQVRIVEGPFATFSGEVEVMSENRKVRVAVTIFGRATPV-ELELVQVEA    |
| 25vib | PRPKTMFEAGEVVRVNDGPFADFGVGVSEVNDYKSRKLVSVISIFGRATPV-ELEFGQVEK   |
| 26xyf | PRPKVLFEPGQVVRVINDGPFNDGVLVEEVNYEKNRLRVAVLIFGRTPV-DLEFGQVQK     |
| 01aae | I-----MARRGKKYIE                                                |
| 04bb  | I-----MSKKGKKYIE                                                |
| 05bs  | L-----MAKKGKKYVE                                                |
| 06cj  | II-----MAKIAKRLKE                                               |
| 07cp  | VAPGQESE-----MTKHGKRIRG                                         |
| 08ct  | VVPGQESE-----MTKHGKRIRG                                         |
| 09der | -----MPKHGKRYRA                                                 |
| 10ec  | A-----MAKLTKRMRV                                                |
| 11hi  | IH-----MAKLTKKMKA                                               |
| 12hp  | II-----VAKKVFKRLEK                                              |
| 13mg  | QIKALYQGTYNKAVVEIELTSKTLINLPLENLSFVEFE-----MKKLSKRMQA           |
| 15mp  | QIKALYQGTINKAVVEVELTTKTLINLPLENLSFIEVEQSH-----MAKLSKMKMI        |
| 17mtu | I-----MSKTSKAYRA                                                |
| 18nm  | IN-----MAKVSRLKA                                                |
| 19pae | V-----MAKLTQRKA                                                 |
| 21rp  | SD-----MSNKDICVKISGSKKIRE                                       |
| 22syn | QN-----MTKKLSKRMQA                                              |
| 23tm  | IE-----MPKHSKRYLE                                               |
| 24tp  | L-----MKRGKKYRA                                                 |
| 25vib | LD-----MAKLTKRMRV                                               |
| 26xyf | S-----MVQ-TKRQKA                                                |
| 01aae | ASKLVDRNKRYTLEEAVDLLKKMEEVLQRRFDETVELAMRLNVDPYADQMVRGSSVVLPH    |
| 04bb  | AFSKVDKNKFYNIEDAILLKKEIK---FAKFDETIDISINLNL---KKNHTVRDTIVLPN    |
| 05bs  | AAKLVDHSAKYDVSEAVLVKKTN---TAKFDTATVEAFRLGVDPSKNHQQIRGAVVLPN     |
| 06cj  | LSQKIDSNEYALSDAIDITIKTLK---SAKFDETVEIALKLNVDPRHADQMVRGSSVVLPA   |
| 07cp  | ILKNYDFSKSYSYSLREAIIDILKQCP---PVRFDQTVDVSVIKLGDIPKSDQQIRGAVVLPN |
| 08ct  | IQETYDLAKSYSYSLGEAIDILKQCP---TVRFDTQTVDVSVIKLGDIPKSDQQIRGAVVLPN |
| 09der | LEGKVDNRNKQYSIDEAAALVKELA---TAKFDETVEVHFRGLGDIPKSDQQNVRGATVLPN  |
| 10ec  | IREKVDATKQYDINEAIALVKELA---TAKFVESVDVAVNLGIDARKSDQNVRGATVLPN    |
| 11hi  | IKAGVDSTKAYEINEAIAVLKQFA---TAKFVESVDVAVNLGIDPRKSDQNVRGATVLPN    |
| 12hp  | LFISKIQNDKAYGVEQGVVVKSALA---SAKFDETVEVALRLGVDPRHADQMVRGAVVLPN   |
| 13mg  | VTKLIDKNLYPIQEAFAELIKKTA---ITKFSVSDVIAVSLNLDTTKAEQQLRGAIAFPF    |
| 15mp  | AVGLVDKTKLYPLQEAVDLVKKTSS---ITKFNGSVDIAVSLNLDTTKAEQQLRGAIAFPF   |
| 17mtu | AAAKVDRTNLYTPLQAALAKETS---STKQDATVEAIAIRLGVDPKADQMVRGTVNLPN     |
| 18nm  | LRSSVEANKLYAIDEAIALVKKAA---TAKFDESVDVSFNLGVDPRKSDQVIRGSSVVLPA   |
| 19pae | IAEKVVAGKQYSFEEAAKLLAELS---TIKFESVDVAVNLGVDPRKSDQVVRGATVLPN     |
| 21rp  | AREKVKSDTLNLTNAVALKKSAS---YVKFDPTLEIVMKLGIDPRHSDQIVRGVNLPA      |
| 22syn | AIKAVDDSKLYSLEAMELLKETA---TAKFDETAEAHRLGIDPKYSDQQIRTTIVSLPK     |
| 23tm  | ARKLVDRTKYDLEAIELVKKTA---TAKFDETIELHIQTGIDYRKPEQHIRTIVLPN       |
| 24tp  | AVARYDRAERFSLDRAVGLLKEVR---YASFDETVEVHVSLRL---KKNTVRDVTVLPN     |
| 25vib | IRAKVDVTKEYDINEAVALLKELA---TAKFVESVDVAVNLGIDARKSDQNVRGATVLPN    |
| 26xyf | IDLAVVPGKAYGIDEAIIKLTAT---KAKFIESVDVAVNLGVDPRKSDQVVRGATVLPN     |

01aae GLGKPIKVVVFA-EGEYAKKAEAGADYVGGDELINKILKEEWDFDVAIATPEMMPKVA  
04bb QFMKPKRILVFA-KGDRADAEARAFGATYVGGDDLINKIKSG-WDFDVVAVATPDMMDVVG  
05bs GTGKTQRVLVFA-KGEKAKEAEAGADYVGGDTYDINKIQQG-WDFDVVAVATPDMMEVVG  
06cj GTGKKVRVAVIA-KDAKAEAKNAGADIVGSDDLVEEIQKG-NMNFVLIATPNLMGLVG  
07cp GTGKTLRILVFA-SGNKVKEAVEAGADFMGSDDLVEKIKSG-WLEFDVAVATPDMMEVVG  
08ct GTGKVLRLVFA-AGDKAAEATEAGADYVGGSDDLVEKIKSG-WVDFDVAVATPDMMEVVG  
09der GTGRSVRVAVIT-KGENVQAAEAGADYVGGDELIERIAGG-FMDFDAVAVATPDMMAQIG  
10ec GTGRSVRVAVFT-QGANAEAAKAGAEVGMEDLADQIKKG-EMNFDVVIASPDAMRVVG  
11hi GTGREVRVAVFT-QGANADAKEAGADLVGMEDLAEQIKKG-EMNFDVVIASPDAMRVVG  
12hp GTGKKVRVAVFA-KDIKQDEAKNAGADYVGGDLAEQIKKG-RIDFDMVIATPDMMAVVG  
13mg SIGKSIRILAIT-DDE--KKASEAGADYVGGDLKIEAIKNG-WLDFDLIITSPKFMGALG  
15mp SVGKPIRILAIT-DDE--KAALAEAGADYVGGIDKINDIKNG-WLDFDLIITSPKFMALG  
17mtu GTGKTARVAVFA-VGEKADAAVAGADYVGGSDDLIERIQCG-WLEFDAIATPDQMAKVG  
18nm GTGKITRVAVFT-QGANAEAAKEAGADYVGFEDLAAEIKAG-NLNFVVIASPDAMRVVG  
19pae GTGKSVRVAVFT-QGPAAEAAAGADYVGMEDLAAEMKGG-DLNYDVVIASPDAMRVVG  
21rp GTGKITRVAVIC-KEEREEAEKAGADLVGSINIIDEIKAG-QINFDVCIATPDMMSIS  
22syn GTGQTVRVAVLA-RGEKVKEATDAGADIAGSEELIEEIQKG-MMDFDLIATPDMMPKIA  
23tm GTGKEVKVLVFA-KGEKAKEALEAGADYVGAEDLVEKIEKEGFLDFVAIATPDMMRIIG  
24tp RFRAEVRVLVFC-KEDRVSEALAGAAYAGGAELYEKVKG-WDFDVVAVASPDMMKDVVG  
25vib GTGRDIRVAVFT-QGANAEAAKAGAEVGMEDLADLVKKG-EMNFDVVIASPDAMRVVG  
26xyf GTGRAVRVAVFVPSGAKAEDALAAGADYVGMDDLAEKMQAG-DLNYDVVIATPDAMRVVG

01aae K-LGRILGPRGLMPSKPTGTVTNVEQAIDAKRGRVEFKVDKAGNVHMPVGKISFEKEK  
04bb R-LGPILGKRGLMPNPKTGTVTNNLKDAINSLKKGRTEFRANKNGVIFSFGKSSMDNEK  
05bs K-IGRVILGPKGLMPNPKTGTVTVEFEKAIGEIKAGKVEYRVDKAGNIHVPVIGKVSFEDEK  
06cj K-VGRILGPKGLMPNPKTGTVTMDVAQAVNNAKSGQVNFVRDQKQNIHAGLGKVSFSKEQ  
07cp K-LGKVLGPRNLMPTKGTGTVTDDAKAISELRKKGIEFKADRAVCNMGVGLSFESSQ  
08ct K-LGKVLGPRNLMPTKAGTGTVTVDVKTVAELRKGKIEFKADRAVCNMGVGLSFESSQ  
09der QKLARLLGPRGLLPNPKSGTGVADVAGMVRGLKAGRIEFRNDKTGVVHAPIGKASFESGN  
10ec Q-LGQVLGPRGLMPSKPTGTVTNVAEAVKNAKAGQVRYRNDKNGIHTTIGKVFDAADK  
11hi Q-LGQVLGPRGLMPSKPTGTVTNVAEAVKNAKAGQVRYRNDKNGIHTTIGKANFSEVQ  
12hp K-VGRILGPKGLMPNPKTGTVTMDIAKAVSNKSGQVNFVRDQKQNVHAPIGKASFPEEK  
13mg K-LGKLLGTGRLMPSKPTGTVTDDVSAIKAYKKGKYEYRDSFGNIHSLGKTDTKTEH  
15mp K-LGKLLGTGRLMPSKPTGTVTDDVPAARAYKKGKYEYRDSFGNIHSLGRVDSASNH  
17mtu R-IARVLGPRGLMPSKPTGTVTADVAKAVADIKGGKINFRVDKQANLHFVIGKASFDEKL  
18nm Q-LGTILGPRGLMPSKPTGTVTNVAEAVKNAKAGQVRYRDTKAGIVHATIGRASFAEAD  
19pae Q-LGQILGPRGLMPSKPTGTVPDVATAVKNKAGQVRYRDTKNGIHTTIGSVGKVDFEPAK  
21rp S-VARILGPKGLMPNPKLTGTVTLDIAKNAKNAKSGQVEYRAEKAGIHTTIGKLSFSDQD  
22syn R-LGKQLGPRGLMPSKPTGTVTADLAAVNEFKAGKLEFRADRTGIVHVMFGKASFADD  
23tm R-LGKILGPRGLMPSKPTGTVTQVEAEAVKEFKGRIEYRDTKNGIHTTIGKVSFSDNEK  
24tp R-LGMVLGRRGLMPSKPTGTVSADLGAACVCELKKGVEFRADKTGVVHLAVGKTTMSAQ  
25vib Q-LGTILGPRGLMPSKPTGTVTNVAEAVKNAKAGQVRYRNDKNGIHTTIGKVTFEADQ  
26xyf K-LGTLLGPRGLMPSKPTGTVSNQNGEAVKNAKAGQVRYRADKAGIHTTIGKVNFDDEA

01aae LIDNLYAADAIVRAKPPGAKGQYIKNMAVSLTMSPSVKLDINEVLKKLQEKAAAEFDK  
04bb IKENYEEFVKEVVKRPSDLKGAFIDSIIYSTMGPSIKVNFVWR-----  
05bs LVENFTMYDTILKAKPAAAGVYVKNVAVSTMTGPGVKVDSSTFNVK-----  
06cj LWDNVSTFVKAINKHKPAAGKRYIKNAALSLTMSPSVKLETQELDMK-----  
07cp IKENIEALSSALIKAKPPAAKQYLVSTFISSTMGPGISIDTRELMA-----  
08ct IKENVEALCAALVKAKPATAGQYLVNFTISSTMGPGVTDTRELIAL-----  
09der LSANYQALISALEGAKPGTAKGVFLRSAYLTTTMTGPSIPLALGGAALA-----  
10ec LKENLEALLVALKKAKPTAGKGVYIKKVSISTMTGAGVAVDQAGLSASVN-----  
11hi LKENLQALLAALNKAAPTAKGFIKKVSIISTMTGAGVAVDQASL-----  
12hp IKENMLELVKTIINRLKPSSAKGKYIRNAALSLTMSPSVSLDAQELMDIK-----  
13mg LVANAMALIDLKSRPSTVKGTIYKNIALTTTMTGPSLKVLPD-----  
15mp LVENALALLDLKSRKPATVKGIYKNIALTTTMTGPSLKVLPD-----  
17mtu LAENYGAAIDVLRKLPSSSKGRYLKKITVSTTTGPGIPVDPSTITRNFAGE-----  
18nm LKENFDALLDAIVKAKPAAAGQYLVKAVSVSTMTGLGIRVDTSSVNN-----  
19pae LQKNVEALLADLRLKLPSSSKGVYKRVTLSTTMTGPGQLQDLASLEA-----  
21rp LLKNLNAFIGAVIKAKPVGLKGNLYKAIYLSSTMGASVQIDLTISI-----  
22syn LLANLALQETIDRNRPSSGAKGRFVRTVSVSSMGPSIPVDINALRDLKFEDN-----  
23tm IKENIIAAIKIMQMPAGVKQGTIKKVVLASTMTGPGIKLNLQSLKE-----  
24tp IVENVDFVLEMDRKKPVVDKAGFVRSISLSSMGPGIWWVHKSEE-----  
25vib LKENLEALLVALKKAKPSSAKGVFVKVSIISTMTGAGVAVDQNTLSAQV-----  
26xyf LKLNQLALLVDLIKIPKPTASKGAYLKQVLSSTMTGPGVMIDQSTLSLK-----

01aae EAYAERERSAASRKTLLKKQELVNSYKERLQKSNGFVFFNFQGDIDAYPLTLRLDIKDL  
04bb -----LFGGNIMSAKINAKKLEMDLLKQFIDSKQNLFFLDYRGLNVAQLTELNRKIEGE  
05bs -----MSSAIE---TKKVVEEIASKLKESKSTIIIV-DYRGLNVSEVTELKQLREA  
06cj -----MTRS---EKVEIIAKLEEGFKASEAIVVC-NYRGLSTKKLEELRNNAEN  
07cp -----MKQ---EKTLLQVEEDKISAAQGFILL-RYLRAFTAAYSREFRNSLSGV  
08ct -----MKE---EKKLLREVEEKITASQGFILL-RYLGFTAHSRSFRNLSGV  
09der -----MANE---KNQQTLSGLKDSLQGIETFYVV-DYQGLTAGQLTQLRKDIREK  
10ec -----MALNLQ---DKQAIVAEVSEVAKGALSAVVA-DSRGVTVDKMTLRLKAGREA  
11hi -----ALNLQ---DKQAIVAEVNEAAKALSASVIA-DSRGVTVKMTLRLKSAREA  
12hp -----MQKHQRQ---HKVELVANLKSQFADAKALLIC-DYKGLSVRKLALRNKARNQ  
13mg -----MVDSSK---NKKQVQVTDVSNLLSQSGFVIF-DYSGMSAVDATLMRKLKFNK  
15mp -----MEAKK---DKAQVADVSHLLTSAGFVIF-DYTSMSAIEATSIRKLLKFN  
17mtu -----MARA---DKATAVADIAAQFKESTATLIT-EYRGLTVANLAEALRRSLTGS  
18nm -----MSLNIE---TKKVAVEEISAAIAQAQLVVA-EYRGISVSSMTLRLANARKE  
19pae -----MAIKLE---DKKAIVAEVNEAAKALSASVVA-DARGVTVGAMTGLRKEAREA  
21rp -----VLR---EKPVAVEDIVNIYKESPSVIT-THYGLTVSQVSSLREELKSK  
22syn -----MGRTRE---NKATVIDSQVQLFQDAQMTVII-DYQGLTVAEITDLNRRLRPL  
23tm -----LLTRQ---QKELIVKEMSEIFKKTSLILFA-DYGLFTVADLTSLRRLREK  
24tp -----MAVRARRLQPAKVAEVESLTRLDEASSYIFT-EYRGLTVEQLTALRALREF  
25vib -----MALNLQ---DKKAIVAEVNEAAKALSASVVA-DSRGVTVAMTNLRKQAREF  
26xyf -----MALNLQ-----MTAMRKQARES

01aae ---KGEIVVGKNTLFYRAFSDTV-LSDHRDIFVG-PTAALFAY-EDPVAVTKKLVFEFLKE  
 04bb ---HGSLKVVKNNIMKMLKEKN-INVVDSCLVG-PTVVVTAL-EEANVIKIFYDFV--  
 05bs ---NVESKVYKNTMTTRRAVEQAE-LNGLNDFLTG-PNAIAFST-EDVVAAPKVLNDF--  
 06cj ---NVKVQIVKNTLANIALNNS---GKTGLVLKDTNIYLWGE-DQLSVSKVAAKFE---  
 07cp ---SAEFEVLKKRIFFKAIEAAG-LEVDCSDTDGHLGVVFSOG-DPVSAAKQVLDFNK--  
 08ct ---SAEFEVLKKRIFFKALETSG-VEMDPEGEGHLLGVVFAVG-DPVSAAKQVLDFNK--  
 09der ---GGQLIVAKNTLLNLALQEGG--RDFDDALKG-PSALVLAQ-EDPAGVAKALSDAAGR  
 10ec ---GVYMRVVRNTLLRRAVEGTP-FECLKDAFVG-PTLIAYSM-EHPGAAARLFKEFA--  
 11hi ---GVTMRVVRNTLLRRAVEGTD-YECLKDTFVG-PTLIAFSN-EHPGARARLFKEFA--  
 12hp ---GIKVQIVKNTLHIAHKET---GYSDDLKETNVFLWGG-DQIALSKLVDFDQK--  
 13mg ---GSKIKIVKNNILRRALKTSN-FEGVDESIVKGIKIAVAVGI-NEILETLKVVDVSV--  
 15mp ---GSKIKIVKNNILRRALKAGK-FEGIDETAIGKLAVAVGV-NEIVETLKAVDGVV--  
 17mtu ---ATYAVAKNTLIKRAASEAG-IEGLDELFGV-PTAIAFVT-GEVDAAKAIKTFA--  
 18nm ---GVYLRVLKNTLARRAVQGST-FAELADQMGV-PLVYAAS--EDAVAAAKVLHQFA--  
 19pae ---GVYKVVVRNTLLKRAVEGTQ-FDVLNDVFKG-PTLIAFSN-EHPGAAARIFREFA--  
 21rp ---EAGFKVVKNTLAKIAAKQTG-LDSITNLFAG-PTAIVYS--KEPVEMARLVVNFA--  
 22syn ---GGTCKIAKNTLVRRALAGQEAWSPEEFLLG-TTALVLK-EDLGAIAKAYKKFQ--  
 23tm YGDGARFRVVKNTLLNALKNAE-YEGYEEFLKGPATVLYVTE-GDPVEAVKIIYNFYKD  
 24tp ---SCVYRVVRNFIANIAFTSLN--MTVGEYLVG-PTAIALVDTEHANGVARVLFDA--  
 25vib ---GVYLVVRNTLARRAVEGTA-YECLKDVFGV-PTLIGFSN-EHPGAAARLFKDF--  
 26xyf ---GVYLVVKNTLARRALGDTE-YAVIKEKLIG-PLLYAFSL-EDPGAAARLKEFS--

01aae TFDKEWEGRIKGLLDYKYITPEQVKELAEPLSKEELIAKLLGLVLMAPVTLAMTLKAVP  
 04bb ---KSSTLVKGGFVLGEFYDEAKVQAYSKLPTKKESISLFAVLKAPVS-----  
 05bs ---KNHEALEIKAGVIEGKSTVEEVKALAEPLPREGLLSMLLSVLKAPVR-----  
 06cj ---ENNDKFEIKTAHIEGEVADVAKVKALAKMPSRNELLAMLQVWNAPIT-----  
 07cp ---QHKDSLVLGARMNDNASLSGAEEVAVAKLPSLKLRRQVVGGLFAAPMSQVVGIMN---  
 08ct ---QHNDSLVFLAGRIDNASLSGREVEAVAKLPSMKELRQVVGGLIAAPMSQVAGIMN---  
 09der ---NDRGIPTVKGGFVEGSKVDVAVVQRLASLSGSKTTLQAEVLVGLSAHLS-----  
 10ec ---KANAKFEVKAFAAFEGELIPASQIDRLATLPTYEAAIARLMTMKKEASAG-----  
 11hi ---KANDKFEIKGAAFEGKIQDV--EFLATLPTYEEAARLMTMKKEAAG-----  
 12hp ---EHKDHVFLKAGLFDKESVSVAHVEAVSKLPSKEELMGMLLSVWTAPAR-----  
 13mg ---KEKELMKFVCGHFDNRIFNSDDLQKIAKLPGRNELVGMFLSVLQAPLR-----  
 15mp ---KAKEAMNFVCGYFDNRFAFNSADLEKIAKLPGRNELVGMFLSVLQAPLR-----  
 17mtu ---KEHKALVKGGMGHPHTVAEVERIADLESREVLAKLAGAMKGNLAKAAGLFNAPA--  
 18nm ---KKDDKIIVKAGSYNGEVMNAQVLAELASIPSREELLSKLLFVMQAPVS-----  
 19pae ---KGQDKFEIKAAAFEGQFLAANQIDVLASLPTYDEAVSGLMSVIQAGTS-----  
 21rp ---NSNDNLKIIGIVDNHILDAYSIKELSKLPSLHELKGIIVGLQAPATKVVGVLQATS--  
 22syn ---KDTKKTRELGGVLEKSLTQADVEAIGDLPSKEQLMGQIAGGINALATKIALGIKEVP--  
 23tm ---KKADLSRLKGGFLGEGKKTAEVENIADLESREVLAKLAGAMKGNLAKAAGLFNAPA--  
 24tp ---KEVPALVVKGAIDLGEVFDASKVEAYSKLPSKEELVSMFLSALNATTV-----  
 25vib ---KENKAFEIKAAAFEGVLTDP--EVLATLPTYDEAARLMTMKKEASAG-----  
 26xyf ---KKHDKLKSXTVSLGGVLYPAGHVDVLSLPTRLQALAMRLVSEPV-----

01aae QKLVVLKAIIEEKS-----KGQG-----MA-KKVATIELMLP  
 04bb ---KLARTLKALADVKN-----MAKKKAIKSWIKLQVP  
 05bs ---NLALAAKAVAEQKE-----EQGA-----VA-KKVVKVVKLQIP  
 06cj ---NFTIGLNALKNKK-----ESE-----MA-KKVVGELIKLQIA  
 07cp ---SVLSGVISCDVQKA-----GKN-----MSVKVVIKIKLQIP  
 08ct ---SVLSGVISCDVQKA-----EKTQE-----MSNKKIIKIKLQIP  
 09der ---NFVGIIEAYREKLEGE---GGSESA-----M-KKVAGIVKLQIP  
 10ec ---KLVRTLAAYRDAK-----EAA-----MA-KKVQAYVKLQVA  
 11hi ---KLARTFAALRDKL-----QEAA-----MA-KKVQAYVKLQVA  
 12hp ---YFVTGLDNLKAK-----EEN-----MA-KKVVGELIKLQIP  
 13mg ---KFLYALQAVRNAK-----MAKKTVTIRIAKINLI  
 15mp ---KFLYALEAVKAAK-----MAKKTITIRIAKINLI  
 17mtu ---SQLARLAAALQEKKACP---GPDSAE-----MAPK-KKVAGLIKLVQIV  
 18nm ---GFARGLAALAEKKA-----GEEAA-----MA-KKIIGYIKLQIP  
 19pae ---KLARTLAAIRDQK-----EAAAA-----MA-KKIQAYIKLVQV  
 21rp ---SSIAIRVIAHAIAKH-----MSKKAIKGYINLIIP  
 22syn ---ASVARGLQAHVDKE-----MA-KKVVALIKLQIP  
 23tm ---RNLVYVNLNAIEKKKS-----E-----MA-KKVAQIKLQIP  
 24tp ---KFVRVLQAVMDKRDGEVSVVSGGDSS-----MAAKKKVVTQIKLQCP  
 25vib ---KLVRTIAAYRDAK-----EAA-----MA-KKVEAYVKLQVA  
 26xyf ---LFARAIKAVADKSETVAVSAPEASQA-MMGSLFLKRYYPEIFQVA-KKVIAYIKLVQV

01aae AQQASPPVPGPALGQHGVMIMEFVKQFNAASRDYEPGTILPVVITVYQDRSFTFIMKTP  
 04bb AAQAAPGAKIGQALGPHGVSGPQFVKEFNERTAKMDPGIIVPVITVYSDKSFSFIVKTP  
 05bs AGKANPAPPVPGPALGQAGVNVMGCFKEFNARTAD-QAGLIIPVEISVYEDRSFTFITKTP  
 06cj ATKANPSPVPGPALGQGVNIMEFCKAFNERTKDMA-GFNIPVVITVYADKSFTFITKTP  
 07cp GGGKANPAPPVPGPALGAGVNMIGFCKEFNAATQD-KPGDLLPVVITVYADKFTFITKTP  
 08ct GGGKANPAPPVPGPALGAGVNMIGFCKEFNAATQD-RPGDLLPVVITVYADKFTSFVMKQS  
 09der AGKATPAPPVPGPALGQGVNIMEFCKAFNAQTAD-KGDALIPVEITVYADRSFTFITKTP  
 10ec AGMANPSPVPGPALGQGVNIMEFCKAFNARTSIEKGLPIPVVITVYADRSFTFITKTP  
 11hi AGMANPSPVPGPALGQGVNIMEFCKAFNARTSIEKGLPIPVVITVYADRSFTFITKTP  
 12hp AGKANPSPVPGPALGQGVNIMEFCKAFNERTKDMG-SFNIPVVITVYQDKSFTFITKTP  
 13mg GGGAKP----GPALASVGINMGFTKQFNEKTD-RQGETIPCIITAFNDKSFTFITKTT  
 15mp GGGAKP----GPALASVGINMGFTKQFNEKTD-KQEMIPVCVITAYNDKSFTFITKTT  
 17mtu AGQANPAPPVPGPALGQHGVMIMEFCKAYNAATEN-QRGNVIPVEITVYEDRSFTFITKTP  
 18nm AGKANPSPVPGPALGQGVNIMEFCKAFNAATQGMPEGLPIPVVITAFADKSFTFITKTP  
 19pae AGQANPSPVPGPALGQHGVMIMEFCKAFNAATQGMPEGLPTPVVITVYSDRSFTFITKST  
 21rp AAGATPAPPVPGPALGQGVNIAAFCKDFNDATNGMEKGVLPPTVITVYEDSSFSFVKVTP  
 22syn AGKANPAPPVPGPALGQHGVMIMEFCKEYNAATAD-KPGMIIPVEISVYEDRSFTFITKTP  
 23tm AGKATPAPPVPGPALGQHGVMIMEFCKEYNAATAD-KAGMILPVVITVYEDKSFTFITKTP  
 24tp AGKATPAPPVPGPALGQHGVSAPQVQVFNDRKSMPEGLPVVITVYSDKSFSFVLKTP  
 25vib AGMANPSPVPGPALGQGVNIMEFCKAFNARTSIEKGLPIPVVITVYSDRSFTFITKTP  
 26xyf AQQASPPVPGPALGQGVNIMDFCKAFNAATQKVEQGLPIPVVITVYSDRSFTFITKTP



01aae PVSYLELLKKAAGVEKGS-SDPKRVKVGKITVKQLEELIAKMKMDMNRDLKAAMRTVAGTA  
 04bb PASILLKKAAGIESGS-KKSNTDKVGTISKEKLMEIARIKMSDLNAKSESAAFKIIAGSA  
 05bs PAAVLLKKAAGIESGS-GEPNRNKVATVKRDVREIAETKMPDLNAADVEAARMVEGTA  
 06cj PATDLIKKAAGISKGT-DNPLKNKVGKLTQAQVLEIVDKKIADLNTKDRDQAAKIIAGSA  
 07cp PVSSLIKKTLNLESGS-KIPNRNKVGKLTQAQVEAIAEQKMKMDIVLLESAKRMVEGTA  
 08ct PVSSLIKKALGLESGS-KIPNRNKVGKLTQAQITVIAEQKMKMDIVLLESARMVEGTA  
 09der PMSYLIRKKAAGIGKGS-STPNKAKVGKLNWDQVLEIAKTKMPDLNAGSVEAAANTVAGTA  
 10ec PAAVLLKKAAGIKSGS-GKPNKDKVGKISRALQEIQAQTKAADMTGADIEAMTRISIEGSA  
 11hi PAAVLLKKAAGIKSGS-GKPNKDKVGKVTLDQVRQIAETKAADMTGATIEKMKSIAGTA  
 12hp PVTDLIKKASGVEKGS-DNPLKNKIATLTHKQVEEIAQLKMEDLNTSMEAAKIVMGSA  
 13mg PVSNIKQAAKLEKGA--KNAKTIVGKISLQQAQEIQAQYKLVLDNANTVEAALKMVLGTA  
 15mp PVSILLKQAAKLEKGA--KNAKTIVGKITMAKAKEIAQYKLVLDNANTVEAALKMVLGTA  
 17mtu PAAKLLKKAAGVAKGS-AEPHKTKVAKVTWDQVREIAETKTKDLNANDVDAAKIIAGTA  
 18nm PASILLKKAAGLQKGS-SNPLTNKVGKLTQAQLEELIAKTKDPDLTAADLDAAVRTIAGSA  
 19pae PAAVLLKKAAGITSGS-ARPNKQKVGTVTRAQLEELIAKTKQADLTAADLDAAVRTIAGSA  
 21rp PASYFLKKYAKITKGSATKKEAMVGKVTIDDCREIAKLKISDLNTKNIEAATKIIICGSA  
 22syn PASVLIKKAAGVEKGS-SEPNKNKVASITREQLREIAQTKLPDLNANDIDAAMNIEGTA  
 23tm PASFLLKKAAGIEKGS-SEPKNKIVGKVTREQLREIAKTKMPDLNANSLEAAMKIIEGTA  
 24tp PAAVLIKKAAGIEKGS-TNSVKQKVARLSLAQLTEIAQVKLPDSALTLDAAKRIIAGTA  
 25vib PASVLLKKAAGIKSGS-ARPNTAKVGTTIDAQIQEIAATKAADMTGADIEAMKRSIAGTA  
 26xyf PASILLKKIVGIGSGS-KRPNTKVGKVTREQLLEIAKTKELDMTAADLDAAVRTIAGSA

01aae KSMGIEIEGWKE-----MKTYRIKPEEVE-RKWVVV  
 04bb RSMGVEVEK-----MNKITNNDTIWPKPT-VE-KKWYVI  
 05bs RSMGIVIED-----MRTTPMANASTIE-RKCLVV  
 06cj RSMGVEIVD-----MTKITKPNVEK-REWIVL  
 07cp RSMGIDVE-----MEKRKDKTTIIVKSSE-TT-KSWYVV  
 08ct RSMGVDVE-----MEKRKDKTTTLAKASDDRN-KAWYVV  
 09der RSMGVTVEGGPNAMAFPDTDVSPRGGPSSPAKSPLRSFVKVTYIPKND-----QNWVVV  
 10ec RSMGLVVED-----MKTFYAKPETVK-RDWYVV  
 11hi RSMGLVVEE-----MKTFYAKPETVK-RDWYVV  
 12hp RSMGVEVVD-----MTKTAKVNDIV-RDWVVL  
 13mg KQMGIEVTD-----MQKTSMLTKEEIAKRNKRWYLV  
 15mp KQMGIEVIE-----MQKTSMLTKEQANKRRQWYIV  
 17mtu RSMGITVE-----VPTYAPKAGDIT-RSWYVI  
 18nm RSMGLDVEGV-----MKTFYAKPEVVK-REWFVI  
 19pae RSMGLNVEGV-----MKTYTAKPETVQ-RDWVFFV  
 21rp ASMGLEVVG-----VKTYSAKPEIE-KKWVVI  
 22syn RNMGITVNS-----MNKTVLPTIDNLD-HKWYVI  
 23tm KSMGIEVVD-----MARYFPVQKTTMIKPEEVE-RKWYVV  
 24tp RSMGVEVERSL-----VRTIFVNERE-AV-RAWHLI  
 25vib RSMGLVVEG-----MKTFYAKPETVK-RDWYVV  
 26xyf RSMGLVVEG-----MTTFYAKNETVQ-RDWYLV

01aae DATGKTLGRLASEIAKILRGKHKPYQPVDVDCGDFVIVINAEKIRVTKKLEQK---KYY  
 04bb DAADRIILGKVAVDVVKILRGKHKAYTTPHQDLGDNVIIINASKVKLTGKYQQK---LYY  
 05bs DAAGKTLGRLSSEVAAILRGKHKPTTYPHVDTDGHDVIIINAEKIELTGKLTDK---YY  
 06cj DAEGKRFGRLLTEVATILRGKHKPCTFPNVDGCDYVIIINASKAVFTGANKAED---KLYH  
 07cp DAAGKTLGRLSSEVAKILRGKHKVTYTPHVMGCDGIVINAEKVRLTGAKKQK---IYR  
 08ct NAEGKTLGRLSSEVAKILRGKHKVTTPHVMGCDGIVINAEKVRLTGAKRAK---VYH  
 09der DASGVPLGRLATLIASRIRGKHPRDFTPNMIQDGFVVVINAQVALTGK-KLDD---KVYT  
 10ec DATGKTLGRLATLARRLRGKHKAEYTPHVDTDGYIIVLNADKVAVTGN-KRTD---KVYV  
 11hi DATGKTLGRLATLARRLRGKHKAEYTPHVDTDGYIIVINADKVAVTGR-KETD---KLYY  
 12hp DAKDKVFGRLITEIAVLLRGKHKRPFTPNVDGCDGFVVVINAANKVFGSG-KLED---KEYF  
 13mg DASGLVLGKLAANLIRGKKNANFTPNQDCGDHLIIINSDQVLTGNKCDNE---FWY  
 15mp DAAGLVLGKLAADLIRGKKNVDFTPNQDCGDYLIINSDQVLTGNKKENE---FWY  
 17mtu DATDVLGRLAVAANLIRGKHKPTFAPNVDGCDGFVIVINADKVAISGD-KLQH---KMVY  
 18nm DAQDKVLGRVAEASRLRGKHKPEYTPHVDTDGYIIVINADKLVGTGA-KFED---KKYF  
 19pae DAAGQTLGRLATEIARRLRGKHKPEYTPHVDTDGYIIVVINAQVVRVTGA-KTTD---KMYV  
 21rp DAKNVVLGRLASRVAIMLRGKHKPSFTPHLDCGDNIINAEHILKTKGLNPKDGKVVY  
 22syn DAEGQRLGRLATEVATILRGKKNPTFTPHMDTDGFVIIINAEKIEVTVGR-KREQ---KLYR  
 23tm DASGKVLGRLATRIAKILMGKHKPNYTPHVDTDGYIIVVINAADKVLTKGKLDQK---VYV  
 24tp DAAGRPLGRVAARVACLLRGKHKASYTPNQEMGDYVVVINAQVFLSGTKPKDK---MYV  
 25vib DAEGKTLGRLASEIASRLRGKHKAEYTPHVDTDGYIIVVINAQKVTVTGN-KAKN---KMYH  
 26xyf DAEGKTLGRLATLARRLLGKTKPVYTHVDTDGYIIVVINAQKVVVTGK-KLTD---KYYH

01aae WHSRYPGGLKERTLKWMLE-NKPEEVIRLAVKRMPLKNRLGHRMLKKLKVYRG-PEHPHQ  
 04bb RHSRYPGGLYSDFRTLSE-RKPCAPLEIAIKGMLPKGPLGRNLFRLNKVFSG-SEHTLK  
 05bs RHTQHPGGLKSRTALEMRT-NYPEKMLEIAIKGMLPKGSLGRQMFKKNLVYRG-SEHPHE  
 06cj RHSGYFGSVKSEKFGDLLE-KNPAKLYKLAVRGMLPKTNLGRAMLKKLKIYAG-SEHPHT  
 07cp YTTGYISGMREIPFENMMA-RKPNYIIIEHAIKGMMPRTLRGKQKLSLIRVKGDSYETFE  
 08ct YTTGYISGMREVPFENMIA-RKPAYVIEHAVKGMPLPKTKLGRQMKSLRVLKGSYAYQE  
 09der RYTGQGLKKTETAREALS-KHPEVIEHAVFGMLPKGRQGRAMHTRKLVYAG-ETHPHS  
 10ec HHTGHIIGGIKQATFEEMIA-RRPERVIEIAVKGMLPKGPLGRAMFRKLKVYAG-NEHNHA  
 11hi WHTGYVGKIKQATFKEMIA-RRPEAVIEIAVKGMLPKGPLGRAMFRKLKVYAG-GEHQHA  
 12hp THSGYFGSTKSLQLEMLE-KAPEKLYHLAVRGMLPKTKLGKAMIKKLKVYRD-DKHPHT  
 13mg HHSQYMGGIKTSGRDMIN-KNSDKLVFNAVKGMLPDNRLSRLITKVHVFKN-DKHNE  
 15mp HHSQYIGGIKVSGRDMLK-KQSDKLVYNAVKGMLPDNRLSRRWITKVHVFKN-DKHNE  
 17mtu HHSQYPGGLHKRTIGELMQ-RHPDRVVEKAILGMLPKNRLSRQIQRLKLVYAG-PEHPHS  
 18nm RSHSGFPGGIYERTFREMGE-QFPGRALEQAVKGMLPKGPLYAMIKKLKVYAG-AEHAHA  
 19pae HHSQFPGGIKSNFEKLIA-KAPERVIETAVKGMLPKNPLGRDMYRKLKVYK-ASHPHT  
 21rp RHTGFPGGIKDTTAKILSGKYPERVIMAVKRMITRNVLGAKQMSNLVYAN-CEHPHM  
 22syn RSHSGRPGMKETFEKLQV-RLPERIVESAVRGMLPKNSLGRKLTFLKLVYAG-PSHPHA  
 23tm HHSQYPGGLKSLTARQMLE-KHPERLILAVKRMPLKNRGRKMLKRLKVYAS-PEHPHQ  
 24tp RSHSGYPGGLKVSFSALVK-RPPEVLEHNAVKGMLPKGPLYAMIKKLKVYAG-SVHPHE  
 25vib RSEFFPGGLKSFSEKLEI-RKPEVLEHNAVKGMLPKGPLYAMIRKLVYAG-AEHNHA  
 26xyf RFTGYVGNLKSLSLQALQ-RHPERVLEIAVKGMLPKGPLYAMIRKLVYK-SKHPHA



```

01aae AQPQPPELVKA-----MIQRQTYLNVADNSGAKKVQVIGIPY-----A
04bb AQNPPIKL-----LIVIMQTYLTIADNTGGKVAQCIVKLG-----GS--
05bs AQKPEVVELRG-----MIQETRLKVDADNSGAREVLTIKVLG-----GS--
06cj AQIAK-EGK-----MIQSFTRLAVADNSGAKELMCIVKLG-----GS--
07cp SQKPILLDI-----MIQESQLKVADNTGAKKVKCFKVLG-----GS--
08ct AIKPIVLDA-----MIQESQLKVADNTGAKKVKCFKVLG-----GS--
09der AQKPQVLKTQPLEVK----MIMPOSRLDVADNSGAREIMCIRVLNSGIGGGLTTGGGG
10ec AQPQPVLDI-----MIQEQTMLNVADNSGARRVMCIVKLG-----GS--
11hi AQPQPVLDI-----MIQEQTMLDVADNSGARSVMCIVKLG-----GS--
12hp AQTSKKDAK-----MIQSFTRLNVADNSGAKEIMCIVKLG-----GS--
13mg AQKPTLLNWS-----MVSFMTRLNVADNTGAKQVGIIKVLG-----AT--
15mp AQKPTLLNWS-----MVSFMTRLNVADNTGAKQVGIIKVLG-----ST--
17mtu AQPQVPVELKQVAQ-----VIQESRLKVADNTGAKEILCIRVLG-----GS--
18nm AQPQKVELK-----MIQMOTILDVADNSGARRVMCIVKLG-----GS--
19pae AQPQVELKI-----MIQTQMLDVADNSGARRVMCIVKLG-----GS--
21rp AQQPTIYDFASENPKNKK--MIQMQSILEVADNSGAKKVMCIVKLG-----GS--
22syn AQPQETLVINTIPAGAN---VIQQQTYLNVADNSGARKLMCLRVL-----GTG--
23tm AQKPEPIEL-----MIQGETYLNVADNSGAKKLRIIRVIG-----G--F
24tp SQNPVPLSC-----MIQVQSRLNVADNSGARLVQCICKVVG-----GS--
25vib AQPQKVLDI-----MIQMOTMLDAADNSGARSVMCIVKLG-----GS--
26xyf AQPQVLDI-----MIQMQSYLVADNSGAKEVMCIVKLG-----GS--

01aae PRKYATLGDVVTVTKKEALPQGNACKGKIYRAIIVRTAKEVRRPDGSYIKFDNACVLLN
04bb KRRYAKIGDIITIVVKQAI PNSSVKKGDVYKAVIVRTSKEVRRKNGTYVRFDDNACVILD
05bs GRKTANIGDVI VCTVKQATPGGVKKGEVVKAVIVRTKSARRSDGSYISFDENACVIR
06cj KRRYATVGDVIVASVKKALPNKGKVKGGVVKAVIVRTKKEIHRDNGSLIRFDENAAVILD
07cp RRRYATVGDVIVCSVRDVEPNSSIKGDVVKAVIVTRRHITRKDGSGLKFDNNSCVIID
08ct RRRYATVGDVIVCSVRDVEPNSSVKKGDVVKAVIVTRNDIHRKDGSTLRFDTNSCVIID
09der NKRYAHVGDIIVASVKDAAPRGAVKAGDVVKAVVVRTSHAIRKADGSTIRFDNAAVIR
10ec HRRYAGVGDI IKITIKEAIPRGKVKKGVLDKAVVVRTKKGVRPDPGSVIRFDGNACVLLN
11hi HRRYAAIGDI IKITVKEAIPRGKVKKGVLDKAVVVRTKKGVRPDPGSVIRFDGNACVILN
12hp HKRYASVGSVIVASVKKAI PNKGKVKRGVVKAVVVRTKKEIQKNGSLVRFDDNAAVILD
13mg YKRYAFLGDVVVVSVKDAIPNGMVKKGGVLRVAVIVRTKKGQQRQDGTGLKFDNACVLIK
15mp RKRYAFLGDVVVVSVKDAIPSGMVKKGGVLRVAVIVRTKKGQQRKDGTHLKFDDNACVLIK
17mtu SRRYAGIGDVI VAVTKDAIPGGNVKRGDVVKAVVVRTKERRRPDGSYIKFDENAAVIR
18nm KRRYASVGDI IKVAVKDAAPRGVRKKGVDVYNAVVRTAKGVRPDPGALIKFDNNAAVLLN
19pae HRRYAGIGDI IKVTVKEAIPRGKVKKGVMTAVVVRTKHGVRRTDGSIIIRFDGNAAVLLN
21rp HHMMAKLGDVIVVSIKEAIPGGKVKKGDVYKGVIVRTKGTGVRSDGSTIKFDKNALVLLN
22syn NCTYGGIGDQIIAVVKDALPNMPIKSDVVRVAVIVRTKQLRRASGMSIRFDNAAVIR
23tm HKKYATVGDIVVCSVREAI PNSDVKKGDVVRVAVIVRTKKEIRNDGTYIRFDNAAVILID
24tp RRRYASVGDIIVVAVKDALPTSVIKKGSVEKAVIVRSKEYRVDGTYIRFDNACVILID
25vib HRRYAHVGDI IKVTVKEAIPRGKVKKGVLDKAVVVRTKGVRRPDPGSVIRFDNACVLLN
26xyf KRRYASIGDI IKVTVKEAIPRGKVKKGEVYDAVVVRTSRGVRPDPGSLIRFDGNAAVLLN

01aae QYGE-PLGTRVLGP IAREVRNKGFTKIASLAPEVV---MVELHDLQP--HWGATKEKKRVG
04bb ANLS-PRGKRVP GPVARELDRANFMKVSLASEVI----MFNLK--PKGASKRRKRVG
05bs DDKS-PRGTRIFGPVARELRENNFMKIVSLAPEVI---MKLHELKP--SEGSRKTRNRVG
06cj NKRE-PIGTRIFGPVGREVRYGGFMKIVSLAPEVL-----MNLTK--AAGSTHKTIRIG
07cp DKGN-PRGTRIFGPVAREIRDGF I KISSLAPEVI---MIKLESFLD-ISERKRKKKLLG
08ct DKGN-PRGTRIFGP IAREIRDGFVKISSLAPEVI---MIKLECLQD-PSPRKRRTKLLG
09der NQGE-PRGTRVFGPVARELDRRFMKIVSLAPEVL---MKLHDLKP--TPGSRKDRKRVG
10ec NNSEQPIGTRIFGPVTRRELSEKFMKII SLAPEVL---MRLNTLSP--AEGSKKAGKRLG
11hi NNTEQPIGTRIFGPVTRRELSEKFMKII SLAPEVL---MRLNTLSP--AEGAKHSKRLG
12hp AKKD-PVGTRIFGPVSVREVRANFMKII SLAPEVV---MIMGLENLKP--AKGSVKKIKRVG
13mg EDKS-PRGTRIFGPVARELREKGYNKILSLAPEVV---MELHQLKSVKSRNHKSKVVG
15mp EDKS-PRGTRIFGPVARELREKGYNKILSLAPEVV---MELNQLKSVKARNHKTTLG
17mtu PDND-PRGTRIFGPVGRELREKRFMKII SLAPEVL---VTLLHDLRP--ARGSKIAIRTVG
18nm NKLE-PLGTRIFGPVTRRELRTERFMKIVSLAPEVL---MFLNTIQP--AVGATHAGRRVG
19pae NKQE-PIGTRIFGPVTRRELTEKFMKIVSLAPEVL---MQLNDLRS-APGARREKRRPG
21rp KQDE-PIGTRVFGPVTRELRAKKYVRIMSLAEVL---MKLNELYN-NLGAKKNKKRIA
22syn AEGN-PRGTRVFGPVARELDRKNFTKIVSLAPEVL---MNLSELSP--KDGAKKRRRVG
23tm KFNA-PRGTRIFGPVARELREKGFMKIVSLAPEVV---MRLDEL RP--TPGAMKKRRRVG
24tp ANGN-PKGKRIFGPVARELRDMFTKIVSLAPEVL---VADFHLIA--PKGANRARRIVG
25vib NNSEQPIGTRIFGPVTRELRNAKFMKIVSLAPEVL---MLLNTLSP--AAGSKHAPKRLG
26xyf NKQE-PIGTRVFGPVTRELRSEKLMKIVSLAPEVLMIMSLRLNDLKP--ALGASSCARVVG

01aae RGIGSGHGKTAGRGHGKQKSRSGDRKMPPYFEGGQTPLYMRIPK-RGFKNP--TRKEYTP
04bb RGPSSGLGKTSGRGKQKQKARN--TSPLRGFEGGQTPLYRRLPK-RGFNS-DYKLEYAI
05bs RGIGSGNGKTAGRGHGKQKQARS-GGVRPFGFEGGQMP LQRLPK-RGFTNI--NRKEYAV
06cj RGQSGMGKTATGKGKQTAR-KGYNEKRGFEGGQPLQRLPK-VGFTSK--IQK-PYV
07cp RGPSSGHGKTSGRGHGKQKQSR-SGYKRRFGYEGGQPL YRRVPT-RGFSHKR-FDKCREE
08ct RGPSSGHGKTSGRGHGKQKQSR-SGYKRRFGYEGGQPL YRRVPT-RGFSHKR-FDKCREE
09der RGPGG-TDKTAGRGHGKQKQSR-GAGKAFFEGGSRSLIARLP-RGFNNV--GTT-YEV
10ec RGIGSGLGKTGGRGHKQKQSR-SGGVRRFGFEGGQMP L YRRLPK-FGFTSR--KAITAE
11hi RGIGSGLGKTGGRGHKQKQSR-TGGVRRFGFEGGQMP L YRRLPK-FGFTSM--KSAVTAE
12hp RGQSGMGKTATGKGKQKQAR-TGYKAKRGFEGGQPLQRLPK-IGFRTK--DSH-IYS
13mg RGHSGLGKTSRGRGKQKQKARKS-GLTRLGFEGGQTP L YRRLPK-YGVANKGILKRWVV
15mp RGHSGLGKTSRGRGKQKQKARKS-GLTRPFGFEGGQTP L YRRLPK-FGNARKGFLKQEVV
17mtu RGDGS-KGKTAGRGTKGTRAR--KQVPVTFEGGQMP IHMRLPKLKGFRNR--FRTEYEI
18nm RGIGSGLGKTGGRGHKQKQSR-SGGFHVGFEGGQMP LQRLPK-RGFSL--TASANAQ
19pae RGIGSGLGKTGGRGHKGLTSR-SGGKVPAGFEGGQPL LHRRLPK-FGFVSL--KAMDRAE
21rp RGIGSGKGTAGRGKQKQKSR--SGVAIKFEGGQTPMIKRLPK-RGFKCI--STKKYNI
22syn RGIAAGQASGSCGMRGKQKQSR-GTGTAKAGFEGGQMP L YRRLPK-LKHFPY--NPKFTV
23tm RGPSSGHGKTSGRGHGKQKQAR-GSGKVHIFEGGQTP LQRLPK-RGFKNI--NKKVYAV
24tp RGSSSGRGTTSRGTGKQKQARAG-HKAYVGFEGGQMP L YRRVPR-RGFSN-AFKKEYAV
25vib RGVSGLGKTGGRGHKQKQSR-SGGKVRPFGFEGGQMP LQRLPK-FGFTSR--KSFVSAE
26xyf RGIGSGLGKTAGRGHGKSFARKGGKIKPFGFEGGQTPMQRLPK-IGFRSR--SVANTAE

```

```

01aae VNVGVIDKLFEDG-----MEITPEVLAS-KGLCDEKDR-VKILG--DGELTKKFVKAHA
04bb VNLGDIDKKFKD-----GQVNYDTLLENKLIKKNKK--IKILS--NGKLTKKVFEVSK
05bs VNLDKLNGFAEG-----TEVTPELLETGVISKLNAG-VKILG--NGKLEKLTVKANK
06cj INVEKITAVKEL-----SEITFESIKSVHKISKSVNK--IKLIG----ASAKDLVSKIK-
07cp ITTGRLAELFQEG-----EAITLDALKAKKAIARQAVR-VKVIL--KGDLEKTFVWQDTA
08ct ITTQRLNEIFDNG-----AEVSLAALKERKVIHRETSR-VKVIL--KGDLEKLVWKDAA
09der VKLSQLQDLEDT-----TFDRDTLEAYRLVRRKNRP-VKLLA--SGEISRAVTVHDA
10ec IRLSDLAKVEGG-----VVDLNTLKAANIIGIQIEF-AKVIL--AGEVTPPTVVRGLR
11hi VRLNELTKVEGN-----VVTLETLKAANILTKDIQF-AKVIL--AGEVKSAVTVRGLR
12hp INVEKNEAIKNL-----EITFSSRLALHHPFLYIEG-VKLIG----KDAKNLASKIK-
13mg LNLNKVAKLN-----LKTVTTRATLIEKKVISKNNLPLKLIG--NTKLTTPHFVQK
15mp LNLNKIAKLK-----LDKINRASLIEKQVISAKSQLPIKLIG--HTKLEKPLHFVHK
17mtu VNVGDINRLFPQG-----GAVGVDDLVAKGAVRKN-AL-VKVLG--DGKLTAKVDVSAHK
18nm LRLSELESIAVN-----EIDILVLKQAGLIASTVSN-VKVIA--SGEISKAVALKGIK
19pae VRTSELAKVEGD-----VVSLLQTKDANLINQHVQR-VKVML--SGEIVRAVTLKGIA
21rp INIYINIEEALTDGRNSTNDIITKEKLE-VGLINNKNL-VKLSICISDDFASPLSKLDA
22syn VNVGKLAGLAPN-----TVVTLESIMEAGIVTSNDGP-LKVLG--NGELAVALT VHAP-
23tm VNVKVLERFEAN-----EEVTPKELIERKIIKDLKDG-VKILG--DGELTKPLVVKHAH
24tp VNVGALEFVYAP-----GETVNRQTLIEKGLVKGRVPF-IKILA--DGELTKSIVVRVDR
25vib VRLSELAKVTGD-----VVDLNALKAANLVTKNIEF-AKIVL--SGEINKAVTVKGLR
26xyf VLSYKLDNLEPG-----EIDFASLRLAKLPSTAKK-AKIVK--KGRLTKVFLVKGIE

01aae FSKSAKEKIEKAGGKWEIIS-----MRHRVKKKHFDRTKEQRLALYRSLARAL
04bb ISKSAESLVIKIGCTIQLV-----MKTGLGFNRLSRKSSHRALLKNMVISF
05bs FSASAKEAVEAAGGTAEVI-----MSYRKLGRTSQAQKAMRLDITD
06cj -----DENISVTGSK-----MRHKHGYRKLGRSSSHRAALLKNLTIAL
07cp VVLS--QGVQNLLGIT-----MQHARKKFRVGRSSSHNRCMLANMLKSL
08ct IVLs--EGVKSLEIAV-----MQHARKKFRVGRSSSHNRCMLANMLKSL
09der ASAAAIKAVEAAGGRVLPVQTTQDDAQAEHRHGKAGRKLNRNNSARVALARAKATL
10ec VTGARAIAEAAAGGKIEE-----MRHRKSGRQLNRNNSHRQAMFRNMAGSL
11hi VTGAKAAIEAAGGSIEE-----MRHRKSGRQLNRNNSHRQAMFRNLASAL
12hp -----DERIKTSGGK-----MRHKHGYRKLGRSSSHRAALLKNLAIAL
13mg ISKNALNAVQTSKGSVKIIT-----MSYINKEGKTTAWRVMTVRQQVSAV
15mp VSKQALKAVENANGSVKLEK-----MSYINPKGKTSARVMTVRQQVSAV
17mtu FSGSARAKITAAGGSATEL-----MPKPTKGPRLGGSSSHQKAILANLATSL
18nm VTGARAIAEAVGGKIEM-----MRHRNGNRKLNRNNSHRRAAMLRNMANSL
19pae ATKGARAIAEAAAGGFED-----MRHRKSGRHLRSTSAHRKAMFQNMAYSL
21rp YSSKAKYLIEKVGQQL-----MRHRKGRRLNVTSSHRQMLANMAVSL
22syn CSKAAQAKIEAAGGSVVAQG-----MRHRCRVPQLGKPADQRKALLRALITEL
23tm FSKSAVEKIESAGGKAEVI-----MRHRVKRHLKGRYGSHRKSLRLNLSREI
24tp VSARAQEKIQAGGSVECEIAQER-----MRHRTGFNPLSCMAAHRRALRNMVTSL
25vib VTGAKAAIEAAGGKIEE-----MRHRKSGRQLNRNNSHRKAMFNSMASSL
26xyf STAGARAMIETTGGSFQE-----MRHQKSGRKNRDTAHRGAMFNSMVASL

01aae IFDERIETTVERAKALRSFIEPLVELA-----KEGTLHARRQALSRLP-----
04bb FKHEKISSTKTLFEVVRFAERLITRA-----KVDTVHNRRELKSFH-----
05bs IINERIETTETRAKELRSVVEKMITL-----KRGDLHARRQAAAYIRN-----
06cj VNSGKIETTLPAKELRGYVERLITRA-----RLGDFNAHRVAFASLQ-----
07cp IHYERIETTLPAKELRRHADKMITLE-----KKNLSAARRIAIGRLMVRYNLTSTK
08ct IHNERIETTLPAKELRRHADKMITLE-----KKNLAARLAVGRLMVRYNLTSTK
09der LREGRIQTTLTKAKELRPFVEQLITTA-----KGGDLHSRRLVAQDIH-----
10ec VRHEIIKTTLPAKELRRVVEPLITLA-----KTDSVANRRALAFARTR-----
11hi VSHEIIKTTLPAKELRRVVEPLITLA-----KVDSVANRRALAFARTR-----
12hp IEHNKIETGIYKAKELRSYIEKLITAA-----RVGDFNAHRVAFAYLQ-----
13mg LSYGKIQTTLKKAQNTQKRLKILITIA-----KVDNFNNRRRAVKKWLLN-----
15mp LAYGKIETTLKKAQNTQKRLDKLITLA-----KVDNFNNRRRVKKWLLN-----
17mtu FEHGRITTTPEKARALRPYAEKLITHA-----KKGALHNRRRELVKLR-----
18nm LTHEAIVTTLPAKELRRVVEPLITL-----KKPSLANRRALAFDRTR-----
19pae FEHEIIKTTLPAKELRRVAEPLITLA-----KEDSVANRRALAFDRTR-----
21rp VIHEQIKTTLPAKELRPYIEALITKA-----KKPDLAVRRSVLSKIK-----
22syn IHRGQIKTTKARAKAVRSEVDRMITLA-----KDGSLAARRRALGYMY-----
23tm VEHGSIVTTTAKAKALKTFMDKLVSKAIEAATDDRARSVHLRRQINAVLG-----
24tp FKFERITTTKPKAAEVRAAERLITRS-----KSDSVHNRQVVARFIW-----
25vib VRHEVIKTTLPAKELRRVVEPLITLA-----KTDSVANRRALAFARTR-----
26xyf FKYQLIKTTLPAKELRRVAEPLITLA-----KVDSVANRRALAFARLR-----

01aae -----DKPAIKKLFEDIAPRFEGRNGGYVRIIKLPYRRRGDGAEMAIIEW
04bb -----DKHILNKLFTKISPVFRQSGGYTRIMIKLGKRY-GDAAEMAIEL
05bs -----EVANEENNQDALQKLFSDIATRYEERQGGYTRIMIKLGRPR-GDAPMAIIE
06cj -----DKNATNKLVTETAPKFKDRNGGYTRIIRKIRRR-GDAAEMAFIEF
07cp EARQAKGGDTSVYNVDRVLVNNLFDLGNFRVERKGGYTRILKLQNRIGDNAQKCIIEF
08ct EARQVKGDL SAYNVDRVIGKLFVLA TRFSSRNGGYTRILKLQNRVGDNAQKCIIEF
09der -----DKDVVRKVMDEVAPKYAERPGGYTRILRVGTTRGDGVTMALIEL
10ec -----DNEIVAKLFNELGPRFASRAGGYTRILKCGFRA-GDNAPMAYIEL
11hi -----NVETVAKLFNELGPRFAQRAGGYTRILKCGFRA-GDNAPMAYIEL
12hp -----NKEATHKLVTETAPKYAQRNGGYTRIQRITFRGDASTLATIEF
13mg -----TNSLDVDQLTNHLFKKVPFLKRNNGYSRVLKLGVRGDSTEMAILQL
15mp -----TNLFDVDQLMDHLFSKVPKYEKTPGGYSRVLKLGPRGDATEMAILQL
17mtu -----DKDVVHTLFAEIGPFFADRGGYTRIIRKIEAR-GDNAPMAVIEL
18nm -----DRDVVKLFGDLGPRFTARNGGYVRVLKYGFRK-GDNAPLALVEL
19pae -----SKAAVGKLFNDLGRYANRPGGYRLILKCGFRA-GDNAPMAYVEL
21rp -----DKRAVEKIIINLGVYKDRPGGYTRIVKSGFYGDLAP IAYIEF
22syn -----DKPTVHALFADAPSRYKDRDGGYTRIIRLRRR-GDAAEMAVIEL
23tm -----DRRLTNKLVDIEAKNYVGRGGYVRVLRIGFRG-GDAAEMSLVQL
24tp -----DKAVLHLKLFADIGPRMREREGGYTRILKGLRQ-GDAAEMVVVLEL
25vib -----DNEVVAKLFNELGPRFAARQGGYTRILKCGFRA-GDKAPMAYIEL
26xyf -----NKEAVGILFSNLGPRVTRPGGYIRLLKCGFRH-GDNAPMAYVEM

```

```

01aae VE-----
04bb VEKPLKVE-----
05bs V-----
06cj VA-----
07cp LAS-----
08ct LA-----
09der V-----
10ec VDRSEKAEAAAE-----
11hi VDRPEVAEATTEE-----
12hp V-----
13mg IDATN-----
15mp TDAKYK-----
17mtu VREKTVTSEANRARRVAAQAQAKAKAAAMPTEESEAKPAEEGDVVGASEPDAKAPPEPPA
18nm VDKP-AAE-----
19pae VDRPVGGEVVEAAE-----
21rp VDRDINAKGNVHQDANEEIKN-----
22syn V-----
23tm VEASSQEG-----
24tp VDYTFEKSCLKRARTDSVPARKGAGKGDASRVSGTVPDQGSQIGIKKKE-----
25vib VDRPEAAVEAAAE-----
26xyf LERP IIAEEVT-----

01aae -----MPRLKTRREKRLRRHKRIRKKVFG---TPERPRLCVYRSLKHFYAQIIDTI
04bb -----MKIKIEAEQRLRRKKRIKDKIGR--GVASRPRTVFKSNRYFYAQVIDDSK
05bs -----MITKTSKNAARLKRHARVRALKSG---TAERPRLNVFRSYKHIYAQIIDVN
06cj -----MRANVLKRKLTLRIKKRIRAKIS---GCENFPRI SVFKSNRTLYIAIDDVK
07cp -----MESSLCKKSLMKRRRALRVKVLK---GSPTKPRLSVVKTNKHIYVQLIDDSI
08ct -----MESSLYKKTSGKARRALRVKALK---GCSLKPRLSVVKTNKHIYVQLIDDVE
09der -----MATATTIRKRLTRRKVRTTTAASGRLLRSVYRSSKHIYAQIIDSR
10ec -----MD-KKSARIRRAATRARRKLQ---ELGATRLVVRHTPRHIYAQVIAPNG
11hi -----MD-KKSARIRRAARARHMMR---EQGVTRLVIHRTPRHIYAQVIAPNG
12hp -----MNAKALYKKALKDRRKLRIKSKLL---GDALRPVSVFRSNRYFYAQIIDVK
13mg -----MTRNDKRRIRHKRIVKKIRLTN-LNNRVVLIVIKSLKNISVQAWDFSK
15mp -----MKTRTEQRLRHKRIVKKIRATN-HDNRVVLMVIKSLNHISVQAWDFSQ
17mtu EAPEN-MAQSVSATRIRSLRRHTRLRKLLSG---TAERPRLVVRHSARHIHVQLVNDLN
18nm -----MD-KHTTLRLRRARKTRARIA---DLKMVRLCVFRSNHIIYAQVISAEG
19pae -----MSVKKETRLRRARKARLKMR---ELETVRLCVYRSSQHIYAQVIAADG
21rp -----MRSAKLKFEKRKSRIRYKISK---TSNRMRLSIFKSCRHIYAQIIDSK
22syn -----MKSTRKSATQRRHRLRLRHLGSG---TSERPRLAVFRSNHIIYAQVIDDVA
23tm -----MIKKESKKEQLRRHRRVRKKVFG---TPERPRLCVFRSNKHIYAQIIDTI
24tp -----MLRKCSDKQRKRMKRKVHIRKRVY---GTAVRPRMTVFRSNRNISVQVIDDDA
25vib -----MD-KKASIRIRRAATRARRKIA---ELGATRLVVRHTPRHIYAQVIAANG
26xyf -----MRRAKSTRVHIR---KLGVPRLSVLRTGRHLYAQIFTADG

01aae GHTLVASSTLDPEFEKITGKRGGKSIDAEVVAEIIARRALEKGIKKVVFDRGGFKYHGK
04bb GHTLIASISTI---EKS LN-LGKN--IDVVKLGEVLAKRLKEKNINLIFDRNGKYHGL
05bs GVTLASASTLKD LN--VESTGD--TSAATKVGLVAKRAAEKGISDVVFDRGGKLYHGR
06cj AVTLAAVD-----GRKLG-VKAN--KEGAKKIAAEFAKTLVKKIEQAVFDRNGYVYHGV
07cp GKTLASVSTLSKLNKSGLTGKN--QEVAKVLGTQIAELGKGNLQDRVVFDRGPFKYHGI
08ct GKTLASISTLAKVAKTSGLTRKN--QDNAKALGIIAELGKGLQVDRVVFDRGAHKYHGV
09der GQTLAAASS---AALKSG--N-K--TDTAAAVGKALAAAAAEKGIVQVVFDRGSYKYHGR
10ec SEVLAAASTVEKAIREQVKYTG--KDAAAAVGKAVERALEKGIDVVFDRSGFYHGR
11hi SEVLAAASTVEKAIREQVKYTG--KDAAAAVGKAVERALEKGVQAVFDRSGFYHGR
12hp QSTITHID-----GRKMG-FKNT--QEDAKKLGALFAELKAGIERAVYDRNGYHGV
13mg NVVLTSSSSLQLKLGK----NGN--KENAKLVGMDIATKLKLNQKQVVFDTGGSKYHGR
15mp NIVLASSSSSLALKLGK----NGN--KDNKLVGQDIADKLKLVKLTNVVFDTGSKYHGR
17mtu GTTVAAASSSEADVRGVP--GDK--KARSVRVGQLIAERAKAAGITVVFDRGGYTYGGR
18nm DKVLAQASTLEAEVRGSLKSGSN--VEAAIIVGKRIAEKAKAGVEKVAFDRSGFYHGR
19pae GKVLASASTLKD LREGA--TGN--IDAANKVQGLVAERAKAAGVTQVAFDRSGFYHGR
21rp SITIAAASTLDDKIKKIKKSHCN--IENAIKVGEIAKKADSAGIKEVVFDRGGYKYHGI
22syn QHTLAAASTLDPDLKKSLSSTAT--QEASAEVGLVAQRAIAKGINQVVFDRGGKLYHGR
23tm GHTLVASSTLVPELREKLQKTWN--VEAAKEVGLLIGKRALEKGIKKVVFDRGGYKYHGR
24tp RSTLASVSTL---EKDFVLLRAN--VSSGLQIGEEIGRRLLEKHIDTVIFDRNGYLYHGV
25vib SEVIAAASTVEKAIREQVKYTG--IDAANKVQGLVAERALEKGVTVAFDRSGFYHGR
26xyf SKVIAAANTLQSQVKDGLKNGKN--SLAATKVGLIAERAKAVGVDRIAFDRSGYLYHGR

01aae IKAFADKCREMGLEFMGVRKLKPV--NGTRHAVLYDFEIEKLVKRGKEWVLLKKNQVEP
04bb IASFATSLREFGINIMGIKTYKPKT--SSLRYKTLSFDDLSK-----GNDP
05bs VKALADAAREAGLKFMIAIKYKPSS--NGRRGMTTSDFAEITD-----K---P
06cj IAALAESLRENGIRLMAIKTYKPYT--PSRRYITGLSSEDITAK-----P---S
07cp VSMVADGAREGGLQF--MFKKFKPV--PGTRQLVLPFDELTTT---GELRGTKSKRSLRP
08ct VAMVADGAREGGLQF--MFKKFKPV--PGTRQLILPSFDELTTQ---GELGSSSRSLRP
09der VKALADAAREGGLDFMAVKKYRPY--PSRRGMTTADFSGLTGK-----R---P
10ec VQALADAAREAGLQFMVAVKCKPTS--PGRRHVVKVVPVPELHKG-----KP
11hi VQTLADAAREAGLQFMVAVKCKPTS--AGRRHVVKVVPVPELHKG-----KP
12hp VAFAESLRENGIALMAIKTYKPYT--PSRRFVMSVLDKSDITAK-----S---S
13mg IAALAEARAKGLNFMIAIKKISRNSNGIHNATVIDFKLLTNSK-----P
15mp IAALAEAREGLNFMPIKISRNSNGIHNSTVIDYKLLTNNK-----KP
17mtu IAALADAAREGLSFMIAIKYKPTT--PGRRGASVDSFAEITRS-----TP
18nm VKALAEAREGLSFMIAIKMKPTS--AGRRGMVVRVTEGLYK-----AP
19pae VKALADAAREGLEFMVAVKCKPTS--AGRRFVVKVVPVPELHKG-----AP
21rp IKALADAAREK--IKFMALKNFNPIT--PSLRELQVQDKTNLWKG-----RP
22syn VKALAEAREAGLNFMIIRNYRPM--PGTRQASVDSFTEITKS-----K---P
23tm VKALADGAREAGLEFMGLKRFKPV--PGRRFMVISDFSDITK-----T---EP
24tp VAAVADGARKAGVKFMALKMYRPM--AGLRFVLDLCRAELT-----ARTP
25vib VAALADSAREAGLKFMIAIKCKPTS--AGRRHVVKVVPVPELHKG-----KP
26xyf IKILADAARDAGLKFMPLIKFKPTS--PGRRSARVVTPIHKG-----SP

```

|       |                                                               |
|-------|---------------------------------------------------------------|
| 01aae | EKSLKWWHRAKGRSRQGNITARHRGGGKKLYRIIDFKR-DKSLVPAKVVSIEYDPFR     |
| 04bb  | LKSLTKGKKFKSGRD-SSGRISIRRRGGGKKRYRLIDFNRRDKFSIPARVASIEYDPNR   |
| 05bs  | EKSLAPLHKKGGGRN-NQKLTVRHQGGGKKRQYRVIDFKR-DKDGIPGRVATVEYDPNR   |
| 06cj  | VRSLLVKPAHAGRN-SYGRITSRHKEAGAKKLYRIIDFKR-KFGIEGKVEAIEYDPYR    |
| 07cp  | NKLSFFKKSSGGRD-NLGHISCRHRGGGAKLYRVVDFKR-NKDGITAKVVTVEYDPNR    |
| 08ct  | NKLSFFKKSSGGRD-NLGHISCRHRGGGVRRHYRVIDFKR-NKDGIEAKVASVEYDPNR   |
| 09der | EKALTEALPKTGGRN-NRGRITSRFISGGGKKLYRIIDFKRRDKSGVNAKVAIEYDPNR   |
| 10ec  | FAPLLEKNSKSGGRN-NNGRITTRHIGGGHKQAYRIVDFKRN-KDGIAPVVERLEYDPNR  |
| 11hi  | YAPLLDTKSKTGGRN-NYGRITTRHIGGGHKQHYRLIDFKRN-KLDIPAVVERLEYDPNR  |
| 12hp  | VKGLLTKLKTAGRN-NNGRITSRHKEGAKKLYRIIDFKRN-KYNIIEGKVAIEYDPYR    |
| 13mg  | EKSLVTLKKHAGRN-NQKQITVRHGGGKKRYRLIDFKRYHYDNLKATVKSIEYDPNR     |
| 15mp  | EKSLVTLKKHGGGRN-NQKQITVRHQGGGRNKKRYRIIDFKRTHYDNIATVKSIEYDPNR  |
| 17mtu | EKSLVRPLHCRGGRN-AHGRITTRHKGKKRAVRMIDFRNDKDGCVNAKVAIEYDPNR     |
| 18nm  | YAPLLEKKNSTAGRN-NNGHITTRHKGKKHHYRVVDFKRN-KDGIAPKVERIEYDPNR    |
| 19pae | YAPLLEKKSNGGRN-NNGRITTRHIGGGHKQHYRLVDFRN-KDGIAPVERVEYDPNR     |
| 21rp  | LKTLTKGMSKTGGRN-QQGRITSRHGGGKKLYRVIDFKR-KIDISAIVERIEYDPNR     |
| 22syn | EKSLTTNRHDQKGRN-NRGVITSRHRGGGKKLYRIIDFKR-NKQNIAPARVAIEYDPNR   |
| 23tm  | EKSLAPLKKTGGRN-HHGRVTVRHRGGGKKRYRIIDFKRYDKAGIPAKVLAIEYDPNR    |
| 24tp  | EKSLTRGKPAKARG-AGGRIVSRHRGGGKKRYRIDFKR-DLHDIPGTVKTIIEYDPNR    |
| 25vib | YAPLLEKNSKNGGRN-NNGRITVRHIGGGHKQHYRLVDFKRT-KDGIAPKVERLEYDPNR  |
| 26xyf | HAALLEPQSKTGGRN-HHGRITVRHIGGGCKQRYRVIDFKRD-KEATIPARVERIEYDPNR |
|       |                                                               |
| 01aae | SARICLLHYADGEKRYIIWPEGLKVGDTVMSISWEDAEAGKPLPEIKPGNAMPLKYIEG   |
| 04bb  | SANIALLVYKDGEKRYIISPKGIVGDVLES-----G-PNAPIKIGNALPLENIPIG      |
| 05bs  | SANIALINYADGEKRYIAPKGIQVGTIMS-----G-PEADIKVGNALPLINIPVG       |
| 06cj  | NCRIALIAKDGKRYILQPRGLSVGDIVAA-----AESGLDIKPGNAMKLKNIPVG       |
| 07cp  | SAYIALLSYEDGEKRYIAPKGIQRGDVVVS-----G-EGSPFKPGCCMTLKSIPLG      |
| 08ct  | SAYIALLNVYDGEKRYIAPKGIKRGDRVIS-----G-EGSPFKPGCCMTLKSIPLG      |
| 09der | SARIALHLYADGEKRYILAPEGLTVGATVNA-----G-PEAEPLKGNALPLRFVPVG     |
| 10ec  | SANIALVLKYDGERRYIAPKGLKAGDQIQS-----G-VDAAIKPGNTLPMRNPVG       |
| 11hi  | SANIALVLKYDGERRYIAPKGLSVGDQIQA-----G-INSPIKVGNSLPMRNPVG       |
| 12hp  | NARIALVVPDGDKRYILQPSGLKVGDSVIA-----AEGGLDIKVGFAAMKLKNIPIG     |
| 13mg  | SCFISLLHYQNGVKTYIISPDKIVGDQVVS-----SDHAIDIKLGYCMPLAFIPEG      |
| 15mp  | SCFVSLITYANGAKSYIISPDKIVGDKILA-----SEHPIDIKPGFSMPLAFIPEG      |
| 17mtu | TARIALHLYDGEKRYIAPNGLSQGDVVES-----G-ANADIKPGNNLPLRNIPAG       |
| 18nm  | TAFIALLCYADGERRYIIAPRGIIQAGAVLVS-----G-AEAAIKVGNLPIRNIPVG     |
| 19pae | TAHIALLYADGERRYIIAPKGVAAGDQLIS-----G-IGAPIKAGNSMPLRNIPVG      |
| 21rp  | TAFIALIKYDDGEYSYIAPKLSIGDRVIS-----S-QAADIKIGNCLPLKSIPIG       |
| 22syn | NARIALLFYTDGEKRYIAPAGLQVGMTVIA-----G-EAPEFIGNTLPLSRIPLG       |
| 23tm  | SARIALLYADGEKRYIAPKGVNVDTLMS-----G-PDAEIRPGNALPLEKIPVG        |
| 24tp  | SVNIALVFIYANGKRYIAPKGLKVGQGVVS-----G-EKVPLEPANALPLGVIPVG      |
| 25vib | SANIALVLKYDGERRYIIAPKGLQAGDVIQS-----G-PDAPIKAGNAMPMRNPVG      |
| 26xyf | TAHIALCYVDGERCYIIAPKGLKEGDKIIS-----G-PNVPKLGNSLPLRNIPVG       |
|       |                                                               |
| 01aae | TIHVNIIEFIPGKGGIARAAGTWAQVLGRSTRKG-YVLVRMPSGEVRMIHERCMATVGRV  |
| 04bb  | RTVHNIELNVKGGQLIRSAAGYAMILAS---DGNVYTVKLSGSEMRILFKKCIATIGEI   |
| 05bs  | TVVHNIELKPGKGGQLVRSAGTSAQVLGK---EGKYVTLRLSGEMRMILSACRASIGQV   |
| 06cj  | TIHVNVLPKPGKGGMIRSAGAYAQLMGK---EEKYVTLRLASGEMRQVLAECMASIGEV   |
| 07cp  | LSVHNIEMRPSSGGKLVRSAAGAAQVIK---SPGYVTLKMPSGEVRMLNEGCRATIGEV   |
| 08ct  | LSVHNVEMRPSGGKLVRSAAGAAQIIAK---TAGYVTLKMPSGEVRMLNEMCRATIGEV   |
| 09der | AVVHALELVPKGAQLARSAGTSVQVQK---ESDYVIVRLPSGELRRVHSECYATIGAV    |
| 10ec  | STVHNVEMKPGKGGQLARSAGTYVQIVAR---DGAYVTLRLSGEMRKVEADCRATIGEV   |
| 11hi  | STVHNVELKPGKGGQLARSAGAYVQIIAR---EGNYVTLRLSGEMRKVLAECVATIGEV   |
| 12hp  | TVVHNIEHMPGAGGQLARSAGMSAQIMGR---ENKYTIIRMPSEMRYILSECMASVGVV   |
| 13mg  | TQVHNIELNPKGGGKLARSAGSYARILGQDE-TGKYVILQLSGETKFLKECRATVGVV    |
| 15mp  | TQVHNIELHPPKGGGKLARSAGSYARILGQDE-TGKYVILQLSGETKFLKECRATVGVV   |
| 17mtu | TLIHAVELRPGGGAQLARSAGSIQLLGK---EASVSLRMPSGEIRRVVDRCRATVGEV    |
| 18nm  | TTIHCIEMKPGKGAQLARSAGASAVLLAK---EGAYAQVRLRSGEVRKINVDCRATIGEV  |
| 19pae | STVHGIELKPGKGAQLARSAGASAVLLAK---EGAYVTLRLSGEMRKVLAECRATIGEV   |
| 21rp  | TTLHNVEMLKGGGKLARSAGTSVELVGK---DSGYAQIKLRSGEFRLVPLDCKATIGSI   |
| 22syn | SEIHNVELVAGRGGMVRSAGAFQVQVAK---EGDYVTLKPSKEVRMVRKECVATLGRV    |
| 23tm  | TLVHNVEFTPGKGGIARAAGTYCQIMAN---EGNYALLRMPSGELRKVHIKCYATVGVV   |
| 24tp  | FTVHNVELTIGKGGIARSAGTRAIVIAAK---DGGYVMLRLPSGEARLVHRCYATIGEL   |
| 25vib | STIHNVELTPGKGAQLARSAGAYAQLVAR---DGAYVTLRLSGEMRKVLSEGRATIGEV   |
| 26xyf | TTVHAVELKPSKGAQMARSAGSSVQLVAR---EGVYATLRLSGEMRRLVLAECRATIGEV  |
|       |                                                               |
| 01aae | GLAEHELVLKLGKAGRARWLGRPHTRGTAMNPVDHPHGGGEGRT-RG-KHPESPWGKTK   |
| 04bb  | GNEDYANISIGKAGSRWLGRRPKVRGAMNPVDHPHGGGEGKTS-G-RHPVSPWGQPTK    |
| 05bs  | GNEQHELINIGKAGSRWKGIRPTVRGSMNPNDHPHGGGEGRAPIGRKSPPSPWGKPTL    |
| 06cj  | GNEEWANVTIGKAGNRHRGIRPQTRGSAMNPVDHPHGGGEGKNSG-RHPVTPWGKPTK    |
| 07cp  | SNADHNLVRDVGKAGRRRWGVRPTVRGTAMNPVDHPHGGGEGRHNG--YIPRTPWGKVTK  |
| 08ct  | SNADHNLVVDGKAGRRRWGVRPTVRGTAMNPVDHPHGGGEGRHNG--YISQTPWGKVTK   |
| 09der | GNAEHKNIVLGKAGSRWLGRPKHQRSAMNPVDHPHGGGEGRTGAG-RVPVTPWGKPTK    |
| 10ec  | GNAEHMLRVLGKAGAAWRGVRPTVRGTAMNPVDHPHGGGEGRN--FGKHVPTPWGVQTK   |
| 11hi  | GNSEHMLRVLGKAGANRWGIRPTVRGTAMNPVDHPHGGGEGRN--FGKHVPTPWGVQTK   |
| 12hp  | GNEDFINVSIKAGNRNRHGIRPQTRGSAMNPVDHPHGGGEGKGTGS-GHPVSPWGTAK    |
| 13mg  | SNLDHNLVVIKAGRSRHKGIRPTVRGSAMNPNDHPHGGGEGRSPVGRDAPRTPWGKRHM   |
| 15mp  | SNLDHNLVVIKAGNRNRHGIRPTVRGSAMNPNDHPHGGGEGRSPVGRDAPRTPWGKRHM   |
| 17mtu | GNAEQANINWGKAGMRWKGRPSVRGVMMNPVDHPHGGGEGKTS-GGRHPVSPWGKPE-    |
| 18nm  | GNEEQSLKKIKAGANRWGIRPTVRGVMNPVDHPHGGGEGRTG-EAREPVSPWGTAK      |
| 19pae | SNSEHSLRSLGKAGATRWGVRPTVRGAMNPVDHPHGGGEGRTS-AGRHPVSPWGLQTK    |
| 21rp  | SNPDQKNINLGKAGNRWLGRPHVRGAMNPVDHPHGGGEGKTS-GGRHPVTPWGFSTK     |
| 22syn | SNAEPRNLKLGKAGRRHLGRPHVRGSMNPNDHPHGGGEGRAPIGRSGVSPWGKPAL      |
| 23tm  | GNEHDHKEVHGKAGRVWLGRPHVRGAMNPVDHPHGGGEGRG-KG--HHPTSPWGLPCR    |
| 24tp  | GNEHDHMTALGKAGARWRGVRPTVRGAMNPVDHPLGGGEGRGK--RNPVTPWGQPCR     |
| 25vib | GNAEHMLRELKAGAAWRGVRPTVRGVMNPVDHPHGGGEGRTS-GGRHPVSPWGVPTK     |
| 26xyf | GNEEHNLRLKLGKAGAKRWLGRPTVRGAAMNPVDHPHGGGEGKSGGQGNHPVTPWGVPTK  |

|       |                                                                |                                   |
|-------|----------------------------------------------------------------|-----------------------------------|
| 01aae | GYKTRRGGKYSQDFIVTRRDGRPL-----                                  |                                   |
| 04bb  | GYKTRKKKRYSDKFIKRRNK-----                                      | MLVNRRTYA                         |
| 05bs  | EFKTRKQKNKSDKFIVRRRNK-----                                     | MQAKA                             |
| 06cj  | GAKTRRK-KASDKLIISRKGK-----                                     | MSK-----A                         |
| 07cp  | GLKTRDK-NKSNKWIVKDRK-----                                      | MFKA                              |
| 08ct  | GLKTRDK-RKSNKWIVKDRK-----                                      | MFKA                              |
| 09der | GLKTRRKRKTSDFIVTRRK-----                                       | MTAPEQTFRNKKQRKQVKLRKPGFA         |
| 10ec  | GKTRSNKR-TDKFIVRRRSK-----                                      | METIA                             |
| 11hi  | GKTRHNKR-TDKYIVRRGK-----                                       | METIA                             |
| 12hp  | GYKTRKK-KASDKLIISRKHK-----                                     | MSK-----A                         |
| 13mg  | GVKTRNMKKHSTNLIIRNRKGEQY-----                                  | MIIFA                             |
| 15mp  | GVKTRNMKKASTNLIIRNRKGEQY-----                                  | MTWWTNSVSSPRPVTSNNTPLTVSMIFA      |
| 17mtu | GRTRNANKS-SNKFIVRRRTGKHSR-----                                 | MTAATKATEYPSAVA                   |
| 18nm  | GYRTRNNKR-TDNMIVRRYSNKG-----                                   | MRVNA                             |
| 19pae | GKTRSNKR-TDNMIVRRRK-----                                       | MEVAA                             |
| 21rp  | GKTRKNKR-TSKFIVKRR-----                                        | MIQKNKNFATA                       |
| 22syn | GAKTRNMKKRSSALIVRRRTK-----                                     | MTKLDTTAEVKA                      |
| 23tm  | GYKTRRGGKRPDKFIVRRNEV-----                                     | MKLQVPRNGLKRSLFHKKKRELLSSLPKIEARA |
| 24tp  | GYKTRKRRVSDRFIVSKRK-----                                       | MTERVTYRA                         |
| 25vib | GYKTRSNKR-TDKYIVRRNK-----                                      | MEATA                             |
| 26xyf | GYKTRKNKR-TQQFIIRGRG-----                                      | MEATA                             |
|       |                                                                |                                   |
| 01aae | -----LFSGKFREKTWETPFTCFVAVIPKKAARIAEKVLKSVIANAE--              | QKGL                              |
| 04bb  | KGKNLPSPPKVRPIADNIRGESYIKAIIVLCSPMNGAKLLEKVVKSAAASNAMYHNKML    |                                   |
| 05bs  | VARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTTPRAASPIIEKVLKSAIANAE-HNYEM  |                                   |
| 06cj  | LIKFIRLSPTKARLIAREVQGMNAELAMASLKFPMPNGAKYIANAISAVANGG----      | F                                 |
| 07cp  | TARYIRVQPRKARLAAGLMRNLVSQEAEEQLGFSQLKAGRCLKKVLNSAVANAE-LHENI   |                                   |
| 08ct  | TARYIRVQPRKARLAAGLMRNLVSQEAEEQLGFSQMKAGRCLKKVLNSAVANAE-SNENI   |                                   |
| 09der | VAKYVRMSPRKVRVLVDVIRGKSVQDAEDLLRFIPRSASEPVAKVLNSAKANAL-HNDEM   |                                   |
| 10ec  | KHRHARSQAQKARLVADLIRGKVSQALDILTYTNKKAALVKKVLESAIANAE-HNDGA     |                                   |
| 11hi  | KHRYARTSAQKARLVADLIRGKVAQALEILTFTNKAALVKKVLESAIANAE-HNDGA      |                                   |
| 12hp  | LLRFVRLSPTKARLIARQIQGMNAELIASLEFTPNKAARVLSKVVASAVANGS----      | L                                 |
| 13mg  | KQYRVHISQKARLVQQLIVGKKINDAQNILNTPKKAAYFLTKLLNSAISNAT-NNHGM     |                                   |
| 15mp  | KQFRVRISQKARLVQQLIVGKKTADAQNILSNTPKKAATLIAKLLNSAIANAT-NNHGM    |                                   |
| 17mtu | KARFVRVSPRKARRVIDLVRGRSVSDALDILRWAPQAASGPVAKVIAASANAQ-NNGGL    |                                   |
| 18nm  | QHKNARISQAQKARLVADLIRGKDVAAQALNILAFSPKKGAELIKKVLESAIANAE-HNNGA |                                   |
| 19pae | KLSGARISQAQKARLVADQIRGKKVGEALNLAFSSKKAELIMKKVLESANAE-HNEGA     |                                   |
| 21rp  | KAKSIRVSPRKLNLVASFIRNMKVSEALVQLTFSPKRIAKIVKDCIQSAIANAE-NNLGL   |                                   |
| 22syn | IARYVRMSPLKVRRLVDQIRGRSYREALIILEFMPYKACEPVKVLRSANAE-HNEGL      |                                   |
| 23tm  | VAKYIRISPRKARAIANTIRGKSVEEAFQILAFSPKKAARIMEKVLKSRVANAE-NNFGL   |                                   |
| 24tp  | KTFLVASPTKVRPVANVVKCPYVRAMALLGHLPHKGARLISKVMKSAASNAIDRDKRL     |                                   |
| 25vib | KHNFARISQKARLVADQIRGKSVDAQLELLTFSNKKAAELVKKVLESAIANAE-HNEGA    |                                   |
| 26xyf | ILRGARISQKARLVAAQVRGLSAESAVNILRFSKKAAELIKKVLESANAE-NNHGS       |                                   |
|       |                                                                |                                   |
| 01aae | DLRLYIKKAVADDGPIKKWIPRAHGRATMVRKRLSHITIVLEEKPEGKEE-----        |                                   |
| 04bb  | SEDMIFVKTMVDDGRRKKIWPRAGRADRLVNRNCHIFVEVDE-----                |                                   |
| 05bs  | DANNLVISQAFVDEGPTLKRFRPRAMGRASQINKRTSHITIVVSEKKEG-----         |                                   |
| 06cj  | EANEVIVKSCRVDAAAVLKRFRPRARGASRIRKPTSHILVEVAKAEVKSEKKTVAKKT     |                                   |
| 07cp  | KRENLSVTEVRVDAGPVYKRSKSRGGRSPILKRTSHLTVIVGEKER-----            |                                   |
| 08ct  | KRENLCVLEVRVDVGPMPFKMKSKSRGGRAPILKRTSHLTVIVGERGQ-----          |                                   |
| 09der | LEDRLFVKEAYVDAGPTLKRILPRARGSANIIKRTSHITIIIAEKGK-----           |                                   |
| 10ec  | DIDDLKVTKIFVDEGSPMKRIMPRAGRADRLKRTSHITIVVSDR-----              |                                   |
| 11hi  | DIDDLKVAKIFVDEGSPMKRIMPRAGRADRLKRTSHITIVVSDR-----              |                                   |
| 12hp  | DAKSALIVSCRVDAGPVLRRSIPRAKGRATAIRKPTSHVFEVAAE--GK--EMK----     |                                   |
| 13mg  | SGDLLYVFECVANQGSPMKRTIARAKSGSVLTKRSSNLVVKLSNPNERNKLLLTQQKEL    |                                   |
| 15mp  | NGDALYVFECVANQGSPMKRTIPRAKGSSNMITKRSSNLVVKLSNPNERNKELIKQKAL    |                                   |
| 17mtu | DPATLVVATVYADQGPTAKRIRPRAQGRAFRIRRTSHITIVVESRPKDKQRSAKSSRAR    |                                   |
| 18nm  | DIDELKVTTIFVDKGPSLKRFPQARAKGRGNRIEKQTCHINVTGN-----             |                                   |
| 19pae | DVDDLKVSTVFNVEGRSLKRIMPRAGRADRLKRSCHITIVKVADK-----             |                                   |
| 21rp  | DIDRLIITKATVKGSVVMKRIMPRAGGRATIRKFFSNLYITVTEKEDN-----          |                                   |
| 22syn | EPADLVVSQAFADGGPSLRFRPRAQGRAYQIRKPTCHITVAAPALADN-----          |                                   |
| 23tm  | SVENLYVSECYVNDGPRMKRIWPRGRGRADIIQKRMSHITVVVRDRSREDEYRKALEELQ   |                                   |
| 24tp  | DEERLFVRDIQIDEGPRLKRLWCRGRGRGDVQLKRMCHITVVVEES--VR-----T       |                                   |
| 25vib | DIDDLRVAKIFVDEGPMKRIMPRAGRADRLKRSCHITIVVADR-----               |                                   |
| 26xyf | NIDALKINTIIVDEGRMLKRFMARAKGRSSRIVKRSSHITIVVGPAPK-----          |                                   |
|       |                                                                |                                   |
| 01aae | -----                                                          |                                   |
| 04bb  | ---KKDIKG-----                                                 |                                   |
| 05bs  | -----                                                          |                                   |
| 06cj  | TTTKAPAKKTTSTKKATVKKES-----                                    |                                   |
| 07cp  | -----                                                          |                                   |
| 08ct  | -----                                                          |                                   |
| 09der | -----                                                          |                                   |
| 10ec  | -----                                                          |                                   |
| 11hi  | -----                                                          |                                   |
| 12hp  | --SSKSHKKNQAEKG-----                                           |                                   |
| 13mg  | --VKKRTMG-HKKEKAKQKQKQ-----                                    |                                   |
| 15mp  | --VKKRVGGQKAKMARQKAVTSVVKAPSKTQGGVQK-----                      |                                   |
| 17mtu | --RTEASKAASKVGATAPAKKAAKAPAKKAPASSGVKTPAKKAPAKKAPAKASETSAA     |                                   |
| 18nm  | -----                                                          |                                   |
| 19pae | -----                                                          |                                   |
| 21rp  | -----                                                          |                                   |
| 22syn | -----                                                          |                                   |
| 23tm  | --KKSISEE-----                                                 |                                   |
| 24tp  | --KDGSKG-----                                                  |                                   |
| 25vib | -----                                                          |                                   |
| 26xyf | -----                                                          |                                   |

```

01aae -----MAAAIKKGDVTLVVRGKEKGKQGVLVVYERVKRRDKQGNPVVY
04bb -----VKTLLKIGDSVKILSGKDRGRIGKIASINRKK-----
05bs -----MHVKKGDKVMVISGKDKGQGTLLAAPPKK-----
06cj -----MAVKLKIKKGSVKVITGDDKGKTKGVLA VYPKT-----
07cp -----MKKQNIIRVGDKVFI LAGNDKGKGVLSLTED-----
08ct -----MKRRSVCVGD TVYVLAGNDKGKQGVLRCLKD-----
09der -----MPRPSAGSHHNDKLHFKKGD TVIVLSGKHKGQTKGVLLALPRD-----
10ec -----MAAKIRRDDEVILVTGKDKGKRGKVKNVLS-----
11hi -----MPAKIRQNDEVILVTGKDKGKRGKVTKVLPN-----
12hp -----MKSEIKNDIVKVIAGDDKGKVA VLAVLPKT-----
13mg -----MQRIKKGDKVVVITGKNKGSGIVLVLTQ-----
15mp -----MQRIKKGDKVVVITGKNKGSGIVLKIMPAR-----
17mtu KGGSD-----MKVHKGD TVLVISGKDKGAKGVLAQY PDR-----
18nm -----MKNIIKGDVVVIA GKDKGKQGVVRLG-----
19pae -----MQKIRRDDEVIA GKDKGKRGKVLKVLAD-----
21rp -----MIKLKVKKGDEVII TGKYKGKGVLVKVPFEE-----
22syn -----MTKTKPAPHRVKMHVKKGDTI QVISGKDKGVGEVLR TIPSH-----
23tm -----MRIKKGDLVEVISGKDKGKRGKVL RVIPKE-----NK-----
24tp -----MGKTVKIRKDDMVLVIA GDRGKRGAVLRVLRDV-----
25vib -----MAAKIRQNDEVILVAGKDKGKGVKTVLAT-----
26xyf -----MASRIKKGQVIVIA GKDKGKQGEIIRIDG-----

01aae RHFVIVEGVRLIKKHVKPIEGVR----EGGI IETEGPIDISNVMLICPNCKNKPTRVGF
04bb -NKVIVESCNMVKKVIKARTPQE----KGRI IDKEAAIDISNMIFV--KGTSSRLGIR
05bs -DRVLVEGVNMVKKSKPTQANP----QGGISNQEAPIHVSNMVPLDPKTEGTVRGYK
06cj -LKVVVEGCKIAKKA IKPSEKNP----NGGFINKEMPDISNVAKVQ--E-----
07cp --KVVVEGVNVRIKNIKRSQKNP----KGKRISIEAPIHISNVRLTI--AGEPAKLSVK
08ct --KVVVEGINVRVKNIKRSQENP----KGRINIEAPLHISNVRLSI--DNQPARLFVK
09der -KQVVVEGVNVITKNVKPSMTNP----QGGQEQLALHASKVALVDPETGKATRVRKQ
10ec -GKVIVEGINLVKKHQKPPALN----QPGGIVEKEAAIQSVNVAIFNAATGKADRVGFR
11hi -GKVFVEGINIITKHEKPPALG----KAGGLVKKEAIDASNVAIFNPNTNKADRVGFR
12hp -SQVVVEGCKVVKKA IKPTDDNP----KGGFIHKEKPMHISNVKKA-----
13mg -NKAIVEGINKVTVHKK-EQVNKRS-KQ-TNP TTKEAPLPLNKLALFDQKAKQQTIGKIK
15mp -QQAIVEGLNKVTRHKKKDDQTTKRAAKQSTGKVQKEAPIFLSKLALFDQKAKQQTIGKIK
17mtu -NRVLVEGVNRKIKHTAISTTQGA--RSGGIVTQEAPIHVSNMVVDSDGKPTRIGYR
18nm -DKVVVEGVNVVKKHQKPNPMRG---IEGGIITKEMPLDISNIAILNPETNKADRVGK
19pae -DRLVVGGVNLIKRHTKPNPMLG---QGGGIVEKEAPLHVSNVAIFNTEKADRVGFK
21rp -NTVVVSGVNLVKKHTKPNKMS----EGGIITQESPIHISNIAHIDPKTGNPTKVAFK
22syn -SQVVVGKVNIRTKHVKPRQEGE---SGQISSYEAPIHSSKVMLYSTKEKIASRICYT
23tm ---VIVENVMVVRHQRPPVQLR---EGGI IEREAPYASKVMVVPACDKRTRVGYR
24tp -DRVLVQGLNMRKKTIRKSAQD---EGGIMEVEAPIHISNMVIMG---KKGPTRVGYR
25vib -GKVIVEGINLVKKHQKPPAMG---VQGGIVEQEAADVSNIAIFNAKTGKADRVGFR
26xyf -HRVVVSNVNI VKRHTKPNQRS---ISGGLIDREAPIHVSNIQVLNPMTKGDRVGFK

01aae IVEEGNVRRKYRYKCKCNENIDL VSEKVIKGG--MPLGLIGEKVGMTRVLLK-DGTAIPV
04bb FENNE---KIRYLKNGQRI-----MLGLIGKVGMTQIFQ-KNGIVVPV
05bs VED-G---KKVRVAKKSGQVLDK-----MTKGLGRKIGMTQVF-AENGDLIPV
06cj -----MEYIVEKIGMSRTIT--NPSIAV
07cp VTEQG---RELWQRPDGTSQLYRLVRGKKG-MRSHISVMGKEGMIHIFD-KDGSVLAC
08ct VREKG---RELWNKHS DGSSSLYRSVRERKG-MRSQSLSIGKKEGMMHVFD-KGNLVLAC
09der IVD-G---KKVRVAVASGKTID-----MKGILGTIGMTQIWK--NDRAIPV
10ec FEDGK---KVRFFKSNSEI I-----MIGLVGKVGMTTRIFT-EDGVSIPV
11hi FEDGK---KVRFFKSNNEI I-----MIGLVGRKVGMTRIFN-EDGVSIPV
12hp -----MEFLVQKIGMSRTID--ANSTPV
13mg YQIDPKTKQKTRVFKKTNNAI-----MDVRGIFGVKVGMSQIFT-EQNECLPI
15mp VYMDPKTNKKTRVFKKSNNTL-----MEIRGIFGVKVGMSQVFT-TNNERLPI
17mtu VDEETG--KRVRI SKRNGKDI-----MARKGILGTKLGMTQVFDE-SNRVVPV
18nm LIENEGKVRVRFFKSNGSIIGA-----MTLGLVGRKVGMTRVFD-EQGVSVPV
19pae VEDGK---KIRVFKSTQKPVQA-----MTIGVGRKCGMTRIFT-EEGVSIPV
21rp FLEDG---SKVRVAKKSGEIIKGVGNVVKV---MRTGIIAQKVGMTSVFN-DKGERISL
22syn VTDG---RKVRMLKKTGEIID-----VSI GILGTKLGMTQIFDQESGISIPV
23tm FTEDG---KKVRYCKKGEI IDKD-----MKM-IIGRKIGMTRVFGV-N-DSVPV
24tp -MENG---KKVRVCRKTGEVL-----MVGLIGQKVGMTQIFD-ARGCVTPV
25vib FEDGK---KVRFFKSNNEI VSN-----MIGLIGRKGMTRVFT-EDGVSIPV
26xyf ILDDG---CKLRIFRSTGEVIGA-----MGFVGRKAGMSRVFL-EDGCSIPV

01aae TVIKFPVNYVVQVKSQNTKDGYNALQIGAYEAKEK-----HLTKPLIGHFKK-HGVPLL
04bb TVIEFPQPNYIIIGKTV D-RDGYSAL IAGSVDLKS-----SKVSKPIKGQYKSLKDIEPK
05bs TVIEAAPNVVLQKKTAE-NDGYEAIQLGFDDKREK-----LSNKP EKHVAK-AETAPK
06cj TLLRVVNKVCVEGGK-----ALVAYPKGKAS-----NKCVAGQKKYNLSAEY
07cp SVIRVEPNVVTQIKTKE-SDGYFSLQIGAEEMNAPAHITITKRVSKPKLGHRLK-AGGRVF
08ct SVISVDANVVQKLTAS-SDGYNAVQMGADV VQAPEKTIEKRFSKALLGHFKK-SGGRVC
09der TVVLAGPCPIVQRKTAQ-TDGYEAVQIGYAPKAER-----KVNKPMQGHFAK-AGVAPT
10ec TVIEVEANRVTVQKDLA-NDGYRAIQVTTGAKKAN-----RVTKPEAGHFAK-AGVEAG
11hi TVIEIEANRVTVQKTL E-NDGYTAVQVTTGSKKAN-----RVTKPEAGHFVK-AGVEAG
12hp TLLKVLQAKVCQLENGK-----ALVAYAMHKKH-----NKAIEGQKKYQLSKEF
13mg TIYVCEANQVAGIKTIA-KDNYNATLLSFQTVDEK-----QLNKPKQGFFSK-LKLEPH
15mp TVIYCEPNQVAGVKTEA-KDKYSATLLSFDTVENK-----KLNKPKQGFFFEK-NNLKPT
17mtu TVVKAGPNVVTRIRTP E-RDGYSAVQLAYGEISPR-----KVNKPLTGQYTA-AGVNPR
18nm TVLDMSANRVTVQVSKD-TDGYTAVQVTFGQKKAN-----RVNKAEGHFAK-AGVEAG
19pae TVIEVEPNRVTVQFKTEE-TDGYRAVQVTA GERRAS-----RVTKAQAGHFAK-ANVAAG
21rp TLVKVDDCQVVGHTLA-KHGYNALVIGVKDKKIS-----KVTKPMKQVFAN-AKIAPK
22syn TVVQAGPCPVTVQKVTQD-TDGYNAIQVGFLPVKEK-----ALSKPELGHLLK-SNTDPM
23tm TVVKAGPCVVVQKKTVE-KDGYNAVQLGF EKAKK-----VNKPLAGHFKK-FGVEPM
24tp TVIRVEHNVVGLKDVE-RFGYSAVILGTGCMKK-----SRISKPYAGQFA--ERIPPV
25vib TVVEVEANRVSVQKLTLE-TDGYAAIQVTTGSKKAN-----RVTKPEAGHFAK-AGVEAG
26xyf TLIEATANRVVQIKTSD-VDGYDAVQVTVGSRRSV-----LVNKPESGHFAK-AKVEAG

```



```

01aae RRLWEFRVDNPEE---FQSQQLKVEDVFKPGDLVDVWGISKGRFAGAMKRWDFAGFPK
04bb RYVIELKGLDG----YDAGDEIKVD-VFKSVKYVDVTGTTKGKGFQAMKRHNFSGGPS
05bs RFVKELRGVEMDA---YEVGQEVKVE-IFSAGEIVDVTGVSXGKGFGGAIKRHGQSRGPM
06cj NRFATLEVAN-----TEAGDLDETP--LNEAKILKVSFNTKGRGYSVGMKRHNFAAGPA
07cp RFLKEVRGSEELNG-VSLGDAFGLE-VFEDVSSVDVRGISKGKGFGQVMKKFGFRGGPG
08ct RVLKEVVVSEEAQS-VSLGDEFGLE-IFDGSVNDVCGISKGKGFGQVMKKFGFRGGPK
09der RILREFRGFAPD-----GDSVND-IFAEGEKIDATGTSKKGKTQGVMKRWNFAAGPA
10ec RGLWEFRL--AEGEE-FTVGQSISVE-LFADVKKVDVTGTSKKGFGAGTVKRWNFRTQDA
11hi RGLWEFR---TEGEE-FTLGQEINVD-IFADVKKVDVTGTSKKGFGQGVKRWNFRTQDA
12hp NHFATLKASQQ-----KELGDLDSA--LETLKRVKASFCTKGRGFAGVMKRWNFQGGPA
13mg KYLREIRKMQG----FELGKKITPQELFKIGEYVDVSLTKGRGFTGAIKR-NFKIGPL
15mp KHLQEIRNMTG-----FEMGQITPQNLQVGEYVDVSAISKGRGFTGAIKRWNFKIGPL
17mtu RYLAELRLDDSDAATEYQVQELTAE-IFADGSYVDVTGTSKKGFGAGTMRKHFRCGQA
18nm RGLIEFALTEEKLAE-LKAGDEITVS-MFEVQGLVDVTGTSKKGFGSGTIKRHNFGAQR
19pae RGVWEFR---LGEEQ-YAAGDQITVD-LFAGQGMVDVTGESKKGFGAGTIKRWNFQGDQ
21rp TKLKEFRISDN---FIDIASILEVD-HFRVGGFVDITATTICKGFGAGSMKRHNFRGLEA
22syn RHLKEYRLTAPN---LQPGDAVTAD-IFQAGDLVDVAGQSMGRGFGYQKRWNFRRGNM
23tm KILREFRVNPDE---YEPGQVIKVD-VFEKGEYVDVTGWTKGRGFAGAMKRWGFSGGPK
24tp RVMREFRGFTLD---VSVQVLDVR-VLESVRYLDVCAISKKGFGQGVKRWGFSGGRS
25vib RGLWEFRL--ENGEE-FAVGSELTVE-LFNEVKKVDVTGTSKKGFGQAVKRWNFRTQM
26xyf RGLWEFRVKTQLGS-YSVGSEVGLS-IFAVGQKVDIQGITKGKGFQGTIKRHNFRMGDA

01aae SHGHR-YHRAVGAIGQR---TDPGRVWKGKRMGPBGHYGAKPVRVQGLFVVDVLPENAILV
04bb SHGSK-FHRHLGGTGQA---TTPARTFKGTMAGRMGNGQTTIQNLEVLIDEERALLV
05bs SHGSR-YHRRPGSMGP---VDPNRVFKGKLLPGRMGGEQITVQNLEIVKDAERNLLI
06cj SHGSR-FHRRHGSIGNR---EWPGRVQPGMKMAGHYGNTKVTIKN-EVVSYDAENKILVV
07cp SHGSG-FHRHAGSIGMR---STPGRCPGSKRPSHMGAEVTVKNLEIVKDLLEKVLVV
08ct SHGSG-FHRHAGSTGMR---STPGRCPGSKRPSHMGCDRVTVKNLEIVKDLDRKVMVLV
09der SHGSKKHRRRPGSIGQR---KTPGRVYKGRMAGHMGMERVTVQNLEIVIRAGENLLV
10ec THGNSLSHRVPGSIGQN---QTPGRVFKGKMGAGQMGNERVTVQSLDVRVDAERNLLV
11hi THGNSLSHRVLGSIGQN---QTPGRVFKGKMGAGHLGAERTVTVQSLVVRVDAERKLLV
12hp AHGSR-FHRRPDSIGNR---EWPGRVQKGRKMAGHYGNELVTCQN-EVLSFDKESMVLVL
13mg GHGAGYPHRFQGSVQAGRGSSAQRFVFKGKMSGHYGHEQVTIQNLFIVGDEINKLVV
15mp GHGAGYPHRFQGSVQAGRGGASQRFVFKGKMSGHYGHEQVTIQNLFIVGDEANMLVLV
17mtu SHGAQAVHRRPDSIGGC---ATPARVFKGTRMAGRMGNDRTVNLVLVHKVDAENGVLV
18nm SHGNSRSRHRVPGSIGMA---QDPGRVFPKGRMAGQYGNATVQKLEIVRVDAERQLLV
19pae THGNSVSHRVPGSIGQC---QTPGRVFKGKMSGHLGAERTVTVQSLIVRVDAERNLLV
21rp SHGVSISHRSHGSGTGR---QDPGRVFKGKMGHMGCNKVTIQNLKIFAVDTNRKLLMI
22syn THGSK-NHRLPGSTGAG---TTPGRVYFGKRMAGQYAGSVTVRRLTVVRVDAERNLLI
23tm SHGSK-FHRELGSVGGH---TEPAKIWKGMKMPGRYGNERTVVRNLQVVDIPDENDLLV
24tp SHGSK-FHREAGSTGC---TSPGRFTKMKMPGRMGAERTVTQNLRIERIDVGLGVVMV
25vib THGNSLSHRAPGSIGQC---QTPGRVFKGKMGHMGAEERTVTQNLRIERVDAERNLLI
26xyf THGNSLSHRAPGSLGQR---QTPGRVFPKMKMSGHMGAERVSVQNLVVRVIDVERCLIAV

01aae KGSVPNGHNGIIVVEQSTIAYRKSQKLQKRYQFIENLAKEESQEVAES-----
04bb KGAVPGAKGSFVVVKSKK-----
05bs KGNVPGAKKSLITVSAVSK-----
06cj KGAVPGYNGAMGKIRIAK-----
07cp KGAIPGARGSIIVKHSSRT-----
08ct KGAIPGFKGSVVVVRSCGVEG-----
09der KGAIPGANGGLVLRSAAKASAAKGK-----
10ec KGAVPGATGSDLVKPAVKA-----
11hi KGSVPGAINGDVIVKPAVKA-----
12hp KGSVAGFSGAYGRIRAV-----
13mg SGAIAGPEGGIVLKTAKKKTGKIKDILAVQTVKAPQLKAPKKQK-----
15mp SGAIAGPEGGVVLRITAKKKGVPVKPIELAVQTEKAPEAKPAKSKKKQAKELAKAQAAN
17mtu KGAVPGRTGGLVHVRSIAK-----RGEK-----
18nm KGAVPGAVNSDVVVRPSVKVGA-----
19pae KGAVPGATGSDVIVRPAAKARG-----
21rp QGSIPGHKNSYLLVKDAIKKAAITIA-----
22syn KGALPGKPTLLNITPAKTVG--RG-----
23tm KGVVPGARGGLVIRSAKAPKK-----
24tp RGAVPGRNKATVFLRTAVKRER-----
25vib KGAVPGATGSGNVIVKPAVKA-----
26xyf RGAIPGASGGDLIRSAKSI-----

01aae -----MKIGDVEV
04bb -----MERKVFSGDKGEIGTINL
05bs -----MPKVALYNQNGSTAGDIEL
06cj -----MSKVVLNDKLEKAGELDL
07cp -----MVLLSKFDGSGNKGVEV
08ct -----MVLLSKFDGSGKESGKFE
09der -----MAQINVIGQNGGR--TIEL
10ec -----MELVLKDAQS----ALTV
11hi -----MELQVVGAN----ALTV
12hp -----MSKAIVLDSHLKEKGSVEL
13mg -----TKVETNQVNPKEIEEETK-----MAKLKVIQFDGSGFGEIQP
15mp QQTVEAKVDTPVVEPKPTVEVKAAPVVEKKGEDK-----MAKLKLIKIDGSEFETEPVK
17mtu -----MAAQEQKTLKIDVKTAPGKVDGAIEL
18nm -----MELKVIDAKGVVSGSLSV
19pae -----MQLNVNGAQ-----AIEV
21rp -----MKTKILSLANEVGEITL
22syn -----MVDCTVKNVQGEVGNASL
23tm -----MAQVDLLNVKGEKVGTLIEI
24tp -----VEKTVSVVEGVALRSVEL
25vib -----MELMVKGAN-----ALTV
26xyf -----MDLTIVGSDN----TLPV

```

```

01aae RD-DVFNVKKV-HVLWEVVKWLAKRRQGTHTSKTRGEVAYSGRKILPQKGTGNARHGE
04bb DD-RVFNIEISH-GSIYNAIKNELSNLRVGTSSKTRSEVRGSSKKPWKQGTGRARVGT
05bs NA-SVFGIEPNE-SVVFDAILMQRASLRQGTHTKVNRSSEVRGGGKPPWRQGTGRARQGS
06cj PS--KY-AEVNP-HNLYLVKSYLASLRANTHTKGRSDVSGGKKPWRQGTGRGARAGS
07cp AD-SLFADEGDGLQIKDYIVAIANKRQWSACTNRSEVSHSTKKPFQKGTGNARQGC
08ct PD-AFFTEGKE--QSVKDYLVAIQANKRQWSACTNRSEVSHSTKKPFQKGTGNARQGC
09der PL---P--EVNS-GVLHEVVTWQLASRRRTASTRTAQVSKTGRKMYGQKGTGNARHGD
10ec SE-TTFGRDFNE-ALVHQVVVYAAAGARQGTTRAQKTRAEVTSGGKKPWRQGTGRARSGS
11hi SE-TTFGREFNE-ALIHQVVVYAAAGARQGTTRAQKTRAEVTSGGKKPWRQGTGRARAGD
12hp PK--RY-ESINS-HNLYLVKHYLSSARANTAKSKNRAEVSGGKKPWAQKGGGRARAGS
13mg AN-HLLKKAVI-QPVFDAILLEQAACRQGTHTSTLTKGEVSGGKKPYKQKHTGKARQGS
15mp TLSPGLIAKELKQ-QPVFDVAVLEQASWRQGTHTSILTKGEVSGGKKPYKQKHTGKARQGS
17mtu PA-ELFDVPANI-ALMHQVVTAAQAAARQGTHTSKTRGEVSGGKKPYRQKGTGRARQGS
18nm SD-ALFAREYNE-ALVHQLVNAYLANARSGNRAQKTRAEVKHSSTKKPWRQGTGRARSGM
19pae SE-RTFGGEFNE-TLVHQAQVVMAGGRQGSKAQKTRSEVSGGKKPWRQGTGRARAGT
21rp NK-DIFAVEFIRDDIKQVLDWQRAKAMFGNHKTITVSEVSGTTKKPFQKGTGNARQGS
22syn TL-RVAKEENAA-HIVHRALVRQNNARQGNASAKTRAEVSGGKKPWRQGTGRARAGS
23tm SD-FVFNIDPNY-DVMWRVYVDMQLSNRRAGTASTKTRGEVSGGKKPWRQKHTGRASHGS
24tp DE-SVFGLSVNR-GVIYYAINSELNKRLLGTACTGRSEVHGSNTKPYKQKGTGRARRGD
25vib SE-TTFGREFNE-ALVHQVVVYAAAGARQGTTRAQKTRSEVSGGKKPWRQGTGRARAGT
26xyf SD-VVFGREFSE-ALVHQIVVAYRNTASGTAQKSRQVSGTTKSKKQKGGG-ARHGA

01aae RGVNIFVGGGVAHGPKP-RDYEYPLPKVRKLGKMLSDKARN-NAIFVDNIDLK--
04bb KRNPMWIGGGIALGPKP-RDYSYRLPKVKKLAFKSVLSRAADENSFKVIENT--NVE-
05bs IIRSPQWRGGGVFGPTP-RSISYKLPKVRRLAISKVLSKVID-NNIIVLEDL--TLD-
06cj TRTNVWVGGAFAVGPTNERNYFQKVNKKQKRLALERALADKAAK-GVLFTADSLAIE---
07cp LASPQFRGGGIVFGPKPKFNQHVIRNKRERKAAIRLLLAQKIQT-NKLTVDVDTVFVDAL
08ct LAAPQFRGGGIVFGPKPKFDQHIRNKKERRAAIRLLLAQKIQT-GKLIYAENSFVSSSL
09der RSVPTFVGGGVAFGPKP-RSYDYTLPRQVRQLGLAMAIASRQEG-GKLVAVDGFDIA---
10ec IKSPIWRSGGVTFAARP-QDHSQKVNKKMYRGALKSILSELVRQ-DRLIVVEKF--SVE-
11hi IKSPIWRSGGTTFAAKP-QDHSQKVNKKMYRGAIKSILSELVRQ-DRLVVVEKF--ELD-
12hp ITSPVFGGGSVSHGATNWRNRYNLKINKQKRLALEYALEEKAQA-NKLFVVEKIAIKGVV
13mg IRNPHYVGGGVFGPKPNRYKLKLNKAYQALTSFAQKLN-NQVIVAEAKLFEQ--
15mp TRNPHFVGGGIVFGPKPNRYSLKLNKAHTAALHTVWSEKLAS-DNTHLVDQNLFNK--
17mtu TRAPQTGGGVVHGPKP-RDYSQRTPKKMIAAALRGALSDRARN-GRHAIITELVEGQ--
18nm TSSPLWRKGGRAFPKPDENFTQKVNKKMYRAGMATILSGLTRD-ERLFAIEAL--TAE-
19pae IIRSPIWRGGGTTFAAKP-RSHEQKLNKKMYRAALRSILAELVRL-DRLVVADF--AVD-
21rp IIRSVQMRGGGISHGPKV-RSHAIKLPKVRKLGILHALSEKYAE-EKLLIDSL--KLD-
22syn IIRSPWRGGGVFGPKP-RDYSQKMNKERRLALRTAASRA---DNMNVVEAFGDQFS-
23tm IIRSPIWRHGGVVHGPKP-RDWSKLLNKKMKLALRSALSVKYRE-NKLLVDLH--KLE-
24tp KKSPLLWGGGTIFGPKP-RDFYALPKVKRLAMKSLSLKAQ--DALTVIEDF--TVE-
25vib IIRSPIWRTGGVTFAAKP-QDHSQKVNKKMYRGAMKSISELSELVRQ-ERLIVVENF--SVE-
26xyf LTAPIFVGGGVAFAAKP-RSFSQKVNKKMYRGAICSIFSELNRQ-GRLLKVVADF--DVE-

01aae -KPK---TKKAIIEFLKNLGE-KEKVLVVIPE-KADV---LYKSFRNLPNVRVLLPE-
04bb -SGK---TKDLALIKNFASF-NGKVILLGN-DDQM---IKRAGKNIRDLKILSFD-
05bs -TAK---TKEMAAIILKGLSV--EKKALIVTAD-ANEA---VALSARNIPGVTVEAN-
06cj -SGK---TKDANAVIKKLGK--DALIVKDL-LDEK---TLAYRNLANCYVVDVT-
07cp -TAPK---TQSLRFLKDCNVE-CRSILFDIHL-DHVEKNENRLSLRNLTAVKGFVYG-
08ct -DAPK---TKEALRFLKECNVE-CRGVLFVDGL-AHVGSNENRLSVRNLSAVRGFTYG-
09der -DAK---TKNFIWAKQNGLDGTEKVLVLT---DDEN---TRRAARNVSWSVLPVA-
10ec -APK---TKLLAQKLDKMALE---DVLIIITGE-LDEN---LFLAARNLHKVDVRDAT-
11hi -APK---TKVLVQKLKDLAVE---DALIITAS-LDEN---LFLAARNLYKVDVRDVQ-
12hp EDNKRKHLTKEANQMFQALEQR---DTLFVCMN-MDEY---TELASNLLKKCLIVDVS-
13mg TNAK---TKKMLTFLKNAKLT-EQKLLFV--I-DTISKP---LLSTNNLKQIV-VKQF-
15mp TEGK---TKVMMQFLKSALL-DKNVLFV--V-NTLNTN---LEQSTSNIKNVQ-VKHL-
17mtu -NPS---TKSARAFASLSTER--KQVLVVIGR-SDEAG---AK-SVRNLPGVHILAPD-
18nm -TPK---TKVFAEQVKNLGL--QVLFVTKQ-LDEN---VYLASRNLPVNLVLEAQ-
19pae -APK---TKGLVAKLDTLGLK--DVLIVTDG-VDEN---LYLAARNLAHVDVRDVQ-
21rp -KPK---TSILVNLLSKFQGG--SFFVIDGNKVDIN---FSLATKNIYNTLIVPQI-
22syn -RQK---TKELATALTRWGAKEPKRVLLILDE-IPEN---VFLSGRNIPYLIKILRAD-
23tm -RPK---TKSLKEILQNLQLS-DKKTLMVLPW-KEEGYM--NVKLSGRNLPDKVIIADN
24tp -SGK---TRDLIQVLRHFAQR-E-RTVFIQNL-DDAL---LKRAGRNIPYLSFLSYN-
25vib -APK---TKALVAKLELELN--DVLIVTGE-VDEN---LFLAARNLYKVDVRDVVT-
26xyf -VSK---TKVFAEKIKSLEVVG--RLLIVSDG-ISEC---LSLSSRNLPDVRVSVQ-

01aae -----GLNVYDVLWANKLVIQKDCLDRIYKKVEA-----
04bb -----KLRVVDLFYAKNLIALESVKNLEFYIK-----
05bs -----GINVLDVNVHEKLLITKAAVEKVEEVL-----
06cj -----EVNAYLVSVFNAVIMEKSVLESITKEG-----
07cp -----ININGYDLASAHNIVISKKALQELVERLVSETKD-----
08ct -----ENISGYDIAARNIVSEKALELLVESLVSTTKD-----
09der -----GVNVYDILRHDLRIDAAALEIVEEAGEEQ-----
10ec -----GIDPVSLIAFDKVMNTADAVKQVEEMLA-----
11hi -----GIDPVSLIAFDKVIITVDVAVKQIEEILA-----
12hp -----ELNAYLLAAFSSSVMEAAAFQHVVDKTEE-----
13mg -----NKVSVRDLLAKTIIIEKAFTKLEERLK-----
15mp -----DKVSVRDMLANALLVEKEVLKALEGKFK-----
17mtu -----QLNTYDVLRADDVVSVEALNAYIAANTTSEEVS-----
18nm -----QVDPVSLRYKVKVIITKDAVAQLEEQV-----
19pae -----GSDPVSLIADKVLVTVSAVKFEELLG-----
21rp -----GANVYDIIRHEYVLLSQEAVSFLEERLR-----
22syn -----NLNIYDVLVADTIVATATALEKIQEYVGE-----VLEFPSSI
23tm PNNSKNGEKAVIDGLNVFDMKYDYLVLTRDMVSKIEEVLGNEAGKALTA-----
24tp -----RLRAHDLFYGRKVLVLETAHVHIAFDYRSKDAADQGT-----
25vib -----GIDPVSLIAFDKVMNTAAAVKQVEEMLA-----
26xyf -----ALDPVALGSDVVVLTVGAVKKIEEWL-----

```



01aae -----MSATETKYVPRLYKKYKEEVVPIQLQKFNYSMPQIPRLQKIVVNMVGVEAVQDI  
04bb -----MNYVPELKKYKDSVIKELVKEFEYKSIQVVPKLEKIVISVGVGEAVRNK  
05bs -----MNRLKEKYNKEIAPALMTKFNYSVMQVPKIEKIVINMGVGDVAQNA  
06cj -----MMRLKEKYNQSIKPALVKEFDIKNPMLIPVIEKVVISVAGELAKDQ  
07cp -----MSRLKKFYTEETIRKSLFEKFGYANKMQIPVLKKIVLSMGLAEAAKDK  
08ct -----MSRLKKLYTEETIRKTLQDKFQYENVMQIPVLKKIVISMGLAEAAKDK  
09der -----MQQLKTKYNDQVRPALMQQFGYSSVMAVPRIEKIVVNEGLGSSKEDS  
10ec -----MAKLHDYKDEVVKLMTFENYSVMQVPRIEKITLNMVGVEAIDK  
11hi -----MAKLHDYRDQVSELKNKFGYKSVMPRIEKITLNMVGVEALTDK  
12hp -----MFGLKFYQSEVRTLKAEQELDKNPMLLPKLEKIVISVAGAGAHAKDM  
13mg -----MNNLEKTYKTELQNLQQLGFSSIMQVPLKIVVNMVGVEAIDRN  
15mp -----MNNLKAHYQKTIKELQKSFASFSSIMQVPRLEKIVINMGVGDARDS  
17mtu -----MTTAQKVQRLKERYSEIRDALRKQFGYGNVMQIPTVTVVVNMVGVEAARDA  
18nm -----MARLREFYKETVPELVKQFGYKSVMEVPRIEKITLNMVGVEAVADK  
19pae -----MARLKEIYRKEIAPKLKEELQANVMEVPRVTKITLNMGLGEAVGDK  
21rp -----MLRFKELYKQKIIESLKKFEFSKNKHEIPKIKKIVINMGVGEAIDS  
22syn YSLTKPRDPVNTMTQRLKTLTYQETILPKLQEEFGYKNIHQVPLKIVVNMVGVEASQNA  
23tm -----MRYEYVPLKQDYKEKIVPALMKELNYKNIHQVPLKIVINMGIGEGSRNY  
24tp -----MTDHSCIPELKVRVYQQIVPDMRDGYSTVMQVPLKIVLSMGLGEALNR  
25vib -----MAKLHDYKSSVVAELTKQFSYTSVMQVPRIEKITLNMVGVEAINDK  
26xyf -----MTRLENMYKKEVVPALIKRFYSNPMAPRLVKITLNMVGVEAATNK

01aae RQLERAVEDLRAITGQQPVITRARKSEAGFKLRKGMPIGCKVTLRKERMWDFLDKLISVA  
04bb KLLDSAVLELAQITGQKAVTKAKKAIAGFKIRGQGEIAGKVTLRGNAMYEFLYKLIHLA  
05bs KAIDSAVEELTFIAGQKPVVTRAKKSIAGFRLREGMPIGAKVTLRGEMYDFLDKLISVS  
06cj KVLQNVADTISLIAGQKAVITAKKSVAGFKVREGFPVGMVTLRKNMYAFLDKLISIA  
07cp NLFQAHLEELTMISGQKPLVTARNISAGFKLREGQGIAGKVTLRGIRMYDFMDRFCNIV  
08ct NLFQAHLEELAVISSQKPLVTRARNISAGFKLREGQGIAGKVTLRGIRMYDFMDRFCNIV  
09der KAIDKAAKELALITLQKPIITKAKKSIKSNFKLRQGMVPGIKVTLRGERMYVFLKELINIG  
10ec KLLDNAADLAAISGQKPLITKARKSVAGFKIRQGYPIGCKVTLRGERMEFFERLITIA  
11hi KLLDNAADLAAISGQKPLVTKARKSVAGFKIRQGYPIGCKVTLRGERMEFFERLITIA  
12hp KIMQNIAGTISLIAGQKAVITAKKSVAGFKIREGMAVGAKVTLRKNRMYNFLEKIVIS  
13mg KFLESALNELHLITGQKPVATKAKNAISTYKLRAGQLIGCKVTLRKNKMSWFLKELIYIA  
15mp KFLESALNELHLISGQKPVATKAKNAISTYKLRAGQLIGCKVTLRGERMAWFLKELIYIA  
17mtu KLINGAVNDLALITGQKPEVRRARKSIAQFKLREGMPVGVVTLRGDRMWEFLDRLTISIA  
18nm KVMHAUSDLEKIAQKQPVVTVARKSIAQFKIRDNYPVGCKVTLRRDQMFELDRDLITIA  
19pae KIENAVADLEKITGQKPVVTVARKSIAQFKIREGWPVGKVTLRSDRMWEFLDRLTISIS  
21rp KVINNALNDLTLISGQKPVVTVARKSIAQFKIREGWPVGKVTLRSDRMWEFLDRLTISIS  
22syn KALESSLTELATITGQKPVVTVARKSIAQFKIREGMPVGMVTLRSEMYAFLDRLINLA  
23tm DLIERHANELAKITGQKPIVTRARKSISNFKIRKGMPIGLKVTLRGARMYSFLYKLINIV  
24tp KLLDASVADLGVISGQHAQKTRARKSIANFKLREGNEIGVMVTLRBSRMYEFLHRLINVA  
25vib KLLDNAASDMAIISGQKPLITKARKSVAGFKIREGYPGCKVTLRGERMWDFLERLISIA  
26xyf KVLNNAVADMAKISGQRPVITKSRISVASFKIRNGWPIGCKTTLRBSRMYEFLDRLINIS

01aae LPRVKDFKGLSPRSFDRGNFYAFGIAEQIVFPEIDYKVDIRGMDIINTTAETDEEAF  
04bb LPRVKDFRGINDAFDGNGNYSFGITEQIIFSEIDYKIERISGLNITIVTTASNDKESK  
05bs LPRVRDFRGVSKKSFDRGNFYTLGKEQLIFPEIDYKVTVRGMDIVITVTTANDDEEAR  
06cj LPRVKDFRGLSRDGFDRGNFYNGLDEQLMFPEVEYDKILRTHGMNISIVTTAQNDKQAQ  
07cp SPRIRDFRGFSNK-GDGRGCVSGLDDQIIFPEINLDRVKRTQGLNITVTTAQTDDECT  
08ct SPRIRDFRGFSCK-GDGRGCVSGLDDQIIFPEVDLDRVKRSQGMNITVTTAQTDACECL  
09der LPRIRDFRGINPNAFDGRGNFYNLGKEQLIFPEIDYDMVDKTRGMDITIVTTAKTDEEAR  
10ec VPRIRDFRGLSAKSFDRGNYSMVGVEQIIFPEIDYKVDVRGGLDITITTTAKNDEEGR  
11hi VPRIRDFRGLSAKSFDRGNYSMVGVEQIIFPEIDYKVDVRGGLDITITTTAKNDEEGQ  
12hp LPRVKDFRGLSRNFDGCGNYTFGINEQLIFPEVYDDIMVSHGMNITVTTSTNDKDEAF  
13mg LPRVRDFRGLSLRSGDKGNFYTGKEQLIFPEIDYDKIRIGFDITIVTSTNDKSEAL  
15mp LPRVRDFRGLSLKSFDRGNFYTGKEQLIFPEIDYDKIRIGFDVTLVTTSTNDKSEAL  
17mtu LPRIRDFRGLSPKQFGVGNFYTFGLAEQAVFHEVDVDKIDVRGMDINVTTSAAATDEGR  
18nm LPRVRDFRGVSGKSFDRGNYSMVGVEQIIFPEIDYDKIDALRGLNITITTTAKTDEEAK  
19pae LPRVRDFRGLNAKSFDRGNYSMVGVEQIIFPEIDYDKIDALRGLDITITTTAKTDEEGR  
21rp LPRVKEFRGFSYKSFDRGNFYTFGLKEQIVFPEINYDKIDITIRGMDITIVTSAKTQDESK  
22syn LPRIRDFRGLSPNSFDRGNYSGLIREQLIFPEIDYDITDQIRGMDVSIITSAQTDEEGR  
23tm LPKVRDFRGLDPNSFDRGNYSFGLSEQLVPELDPDEVRIQGMNITIVTTAKTDQEVVR  
24tp LPRVKDFRGVSPKQFGVGNYSMVGITEQIIFPEIDYDKIERISGLNVNVVTSAAQTDEAR  
25vib LPRVRDFRGVNGKSFDRGNYSMVGVEQIIFPEIDYKVDVRGGLDITITTTAGTDEEGR  
26xyf LPCVRDFRGIIPRSFDRGNFNMVGVEQVVFPEIDFADVAIRGMDIAITVTTANSDAEAK

01aae WLLSLLGLPIRSM-----MSRLAKKPIPY-PENVKVN  
04bb ALLLKFGMPFSN-----MSRIGRLPIKI-PDAVKVD  
05bs ELLTQVGMPPFK-----MSRVGKKLEI-PSDVTVT  
06cj KLELIGVPFTKGK-----MSRIGKPIAI-PAGVEVK  
07cp TLLELMGLRPFKKAQ-----MSRKAREPILL-PQGVES  
08ct TLLECMGLRPFKKAQ-----MSRKARDPIVL-PQGVES  
09der ALLQSMGLPFRKQ-MPQRRRLHHLDLTSAPRDLRNKEDNMSRIGKQPIAV-PSGVTVN  
10ec ALLAAFDFFPRK-----MSRVAKAPVVV-PAGVDVK  
11hi ALLAAFNFPFRK-----MSRVAKAPVNI-PAGVEVK  
12hp KLELLGLPFAKVR-----MSRIGKRIE-I-PSSVQAS  
13mg ALLRALKMPFVKE-----MSKIGNRSKIDPSKVSML  
15mp ALLRALNPLVKG-----MSKIGNRTITLPAKVNLN  
17mtu ALLRALGFPFKEN-----MSRIGKQPIPV-PAGVDVT  
18nm ALLSLFKFPFKG-----MSRVAKNPVTV-PAGVEVK  
19pae ALLRAFKFPFRN-----MSRVAKNPVKL-PAGVEIK  
21rp FLLSGFNLPFYN-----MSRVGKLPIIT-PEGVKVG  
22syn ALLKALGMPFRS-----MSRIGKRPIL-PAKVSVD  
23tm RLLELFGMPPFRG-----MSRLAKKPIVL-PQGVTV  
24tp TLLTKLGMPPFRK-----VSRIGKVPVS-PGGVHVR  
25vib ALLAAFNFPFRK-----MSRVAKAPVAI-PAGVEVK  
26xyf ALLDAFNFPFRN-----MSRVAKPISV-PKGVES

01aae YVEK-EHKIVVEGPKGKLELVNHP--DIKVTNQQ-ERWIKLD-----RPTDRSFHKAH  
04bb VK---DNLVIVEGIRGLVQDIKD--SINVKVENG-SVIVDRVLN-----DKKAKAYH  
05bs LN--DNNTVAVKGPKGELTRTFHP--DMEIKVEDN-VLTVARPSD-----QKEHRALH  
06cj LE---GNLLKFKK--GNLAKELDTKANVNVEIKDN-NILFSPKGE-----DRQSRAIW  
07cp IQ---DDKIIIVGPKGSLTQKSVK--EVEITLKDN-SIFVHAAPHVVD-----RPSCMQ  
08ct IQ---NDEISVKGPKGSLTQVLAK--EVEIAVKGN-EVFVTPAAHVVD-----RPGRIQ  
09der AQ---DGVFKVKGPKGELTVPYNT--ELTVRQDGD-QLLVERPSD-----AQKHRALH  
10ec IN---GQVITIKGKNGELTRTLND--AVEVKHADN-TLTFGPRDG-----YADGWAQA  
11hi LD---GQLLTVKGKNGELSRKIEH--SVEVKQDNG-QFTFTPREG-----FVEANAQS  
12hp VE---GSKLLFKN--SKEKHELETHNRVKITLENN-QLSFQPVGE-----DAQSRAIW  
13mg QT---TLLTIKGPLGENTIKLPKNLPLKFVNVEND-TIKVTNNNN-----LKQTKILH  
15mp FQ---KDHIIVKGPLGQIELKLPPNLPLKFELKDN-NLQITRNN-----LKQSKIFH  
17mtu IE---GQISIVKGPGLTGLTVAE--PIKVARNDGAIIVTRPDD-----ERRNRSIH  
18nm FG---AEALVIKGNKNGELSPPLHS--DVAIEFNDG-KLTFVANNS-----SKQANAMS  
19pae LA---GQLLSIKGAKGALELVKHP--SVEVIQDSG-ELRFAARNG-----DQTRAMA  
21rp L--N-DLEVKISGPKGELSKTFKG--NIAIMEEN-KLVVKPL-----AVKNARAMW  
22syn IQ---GSHLSVKGPKGSLERQLPE--KVIVAQEGE-TITVTRQDESR-----TARERH  
23tm I--K-DNVVKVKGPKGELSQEFLP--YVKIEVEGN-EGWVRPNEEQIRKSDWRKVKMFQ  
24tp VS---SGVVEVEGPKGVLSCAFLP--VVTVRVEQE-YVIVARCD-----SKRARACH  
25vib LN---GQETIKGAKGELTRVFHN--GVVIAQEDN-QLTFGPREG-----VANAWAQA  
26xyf VR---SDMLTVKGVKGLTTFPKSD--NMNVVVDGD-ILTL SAND-----PSHVLSA

01aae GTMAALVKNMIKGVTEGFTVELEIHLGLYRAQLKGNVLEHLGKSHPDYIPDPVKIEV  
04bb GLYRSLIFNMVKGVTGFSKSLTINGIGYRVEQQNSLFLSLGYSTQFEYVIPDGISVKL  
05bs GTTRSLGNGMVEGVSKEFERGELVGVGYRASKGNKLVNVGYSHPVIEPEEGIEIEV  
06cj GYRALAYNIVVGLTQGSFKTLEINGVGYKAALKGVLELSLGFSPHINYDIPGEGIEIVV  
07cp GLYVALISNMVQGVHLGFELKLEMIGVGFRAVQGAFLDLISIGVSHPTKIPSTLQVSV  
08ct GLYVALIANMVKGVHTGFELKLEMIGVGFRAAVQGSLLDLISIGVSHPTKIPSTLQVSV  
09der GTARTRLVANAVKGVSDGYTINLELRGVGFRAKTGKALEMNIYSHPVIEPPAGVTFAV  
10ec GTARALLNSMIVGTGFTKKLVGVGYRAAVKGNVNLGLGFSHPVDHQLPAGITAEAC  
11hi GTARALVNAMVIGTGTGFTKKLVGVGYRAQLKGNALSLGYSHPVHTLPVGITAEAC  
12hp GTYGALANNIVIGLSTGFSKTLVNGVGYKVALGNKTLDLGLGFSHPVKYPIPAGIEMVV  
13mg GTFNALVNNAVIGTVKGFEKLLIVGVGYRANVEGQFLNLQGLYSHPIKELIPNQLTVKV  
15mp GTYNALITNAIIGVTQGFELKRLVGVGYRANVEGETLNLQGLYSHPIKELIPKGLTVKV  
17mtu GLSRTLVSNLVTGVTGYTITKMEIFGVGYRVLKGNLALFGLYSHPVIEAPEGITFAV  
18nm GTARALVSNMVKGVSEGFELKQLIGVGYRAAQKILNLGLGFSHPVYEMPEGVSVQT  
19pae GTTRALVNMMVGVSGQFERKQLVGVGYKAQAGQVLSLGLGFSHPVDYELPAGIVAE  
21rp GTARSIIICNMITGVKEGFKLLEINGVGYRAMVKGYLNLMLAKSHNTKIEIPSNIKIDL  
22syn GLVRTL VANMVDGVAQGFERRLEIQGVGYRAQAGNKLTLNVGYSKPVEMTPQGIIEVKV  
23tm GTYWSLIRNMVGVTEGYKKELEIVGIGYRAQLQGNLVMNLGYAHPVYIEPSDVKIEV  
24tp GLYRKLLSNMVGVSSEGFELKTLVITGIGYRAEVQGRVLVMAVGYSNDFTVLIPSGIEVRV  
25vib GTARALVKNMVGVTGFTKKLVKGVGYRAAMKGNVGLTLGFSHPVEHELPAQVKAEC  
26xyf GTVTRAILSNMIGVSGFERKLELVGVGYRASMQENLNLGLGFSHPVLFVPEGITLLT

01aae KG-NEIHIGIDKQVRGVQAAEIRSFRRKPDYKGGKIRYKGEQLKLPKPG----KAVGKK  
04bb DGNTKISVEGIDFKVGVQAAEIRSLKPEPYKGGKIKYDNEVIRRVKVG----KSGVKK  
05bs PSQTKVVVKGTDKERVGAIAANIRAVRSPPEPYKGGKIRYEGEVRRKKEG----KSAK--  
06cj DKNT-IAVKGSDDKQVGVQAAQIREFRPPEPYKGGKIRYENYVRRKAGKTS--KK---  
07cp EKNTLISVKGDLKQVGEFAASIRAKRPPPEPYKGGKIRYENYVRRKAGKAA--KTGKK-  
08ct EKNTLISIKGINKQVGEFAACVRAKRPPPEPYKGGKIRYENYVRRKAGKAA--KTGKK-  
09der PEPTRIDVSGIDKQVGVQAAANVRKVRKPDAYHGKGVRFVGEQIALKAG----KAGATG  
10ec PTQTEI VLKGADKQVIGQVAADIRAYRRPEPYKGGKIRYADEVVRTRKEA----KKK---  
11hi PSQTEI VLKGADKQVIGQVAADIRAYRRPEPYKGGKIRYADEVVRTRKEA----KKK---  
12hp EKNT-IITKGSDDKQVGVQAAEIRSFRRPPEPYKGGKIRYENYVRRKAGKTA--KK---  
13mg EKNTETISIGIKELVQGFATEIRKWRKPEPYKGGKIRYENYVRRKAGKTAEGKK---  
15mp EKNTETISIGIKELVQGFATEVRKWRKPEPYKGGKIRYENYVRRKAGKTAEGKK---  
17mtu QAPTFTVSGIDKQVGVQAAANIRRLRRPDPYKGGKIRYEGEVRRKKEG----KTGK--  
18nm PSQTEI VLTGSDKQVGVQAAEIRAFRAPEPYKGGKIRYEGEVVRMKEA----KKK---  
19pae PSQTDILIKGIDKQVGVQAAEIRDFRPPPEPYKGGKIRYADEVVRTRKEA----KKK---  
21rp PKQNIILLEGIDKEKLGQFASIIKQRPPEPYKGGKIRYENYVRRKAGKTS--KKN---  
22syn ENNTQVIVSGIDKELLGNTAAKIRAVRPPPEPYKGGKIRYEGEVRRKAG----KTGKK-  
23tm PAPNRIIVSGIDKQVRGVQAAEIRAFRPPNYTGGKIRYEGEVVRKQEG----KKA---  
24tp ESSTRVIVSGVSKERVGEFAAQLRLRLPEAYKGGKIRYDYETIVRKVG----KSGVK-  
25vib PSQTEI VLTGCDKQVGVQAAADIRSYRAPEPYKGGKIRYADENVRKKEA----KKK---  
26xyf PSQTEVVVQIDKQVRGVQAAKIRSFRRPPEPYKGGKIRYAEATMRKEA----KKA---

01aae -----MATLTIDEIVAEIAKNMSVLEVAELVKRLEEEFGVSAAA--MVAAPAAGAAAGAP  
04bb -----MALNKEDILTWLEGAKTMEVVDLVTAEIEEFVGTAAVAGGVGAAVSGSAD---  
05bs -----MA-LNIEEIIASVKEATVLELNDLVKAEIEEFVGTAAAPVAVAG-GAAAGG---  
06cj -----MAISKEDVLEYISNLSVLESELVKFEKFGVSAAPVMVAGG--AVAGGA-V  
07cp -----MTTESLETLVEKLSNLTVLELSQLKKLLEEKWDVTASAPVVAVA--AGGGGE--A  
08ct -----VTTESLETLVEQLSGLTVLELSQLKKLLEEKWDVTAAAPVVAVAG-AAAAGD--A  
09der GKGGK-MAYDKQALIDQLGQLTIMELADLIDGLTEGWVTAAPVAVSGGG--AGAASP---  
10ec -----MS-ITKDQII EAAVAMSVMVDVELISAMEEKFGVSAAPVVAVA--AGPVEA---  
11hi -----MS-LTNEQII EAIASKTVTEIVELIAAMEEKFGVSAAPVVAVAAGGAAA---  
12hp -----MAISKEEVLEYISLVSLESELVKFEEKFGVSAATPTVVAAG--AVAGGA--A  
13mg -----MGKLDKKQLIESLKEMTIVEIDEIIKAVEEAFGVTATPIVAA--AAGAT---  
15mp -----MAKLDKNQLIESLKEMTIVEIDEIIKAVEEAFGVSATPVVAA--AVGGT---  
17mtu -----MAKLDSTDELLDAFKEMTLLELSDFVKFEETFEVTAAPVVAAGAAAPAGAAV-  
18nm -----MAITKEDILEAVGSLTMELNDLVKAEFEKFGVSAAPVVAVAG--AGAG---A  
19pae -----MA-LTNEDII NAVSEMVMQVVELIKAMEEKFGVTAAAATVAA-AGPAAAA---  
21rp -----MADLAKIEEQLSSLTLMQAAELVKMLEEKGVSAAPITVAS-AGVAAPL--A  
22syn -----MSAATDQILEQLKSLLEAESELVKQIEEAFGVSAAPVGGVMMAAAAP--A  
23tm -----MTIDEII EAIKLTVSELAEVLKLEKDFGVTAAAPVVAAGAAAPVAGAAAG---  
24tp -----MAALSNEQII EAIKGTILLESELKIAVEEFGVTAAPVPAVAGGGGAGS---  
25vib -----MS-ITNEQILDAIEMSVMQVVELISAMEEKFGVSAAPVVAAG--PAAAAA---  
26xyf -----MS-LTNEKIVEAIAEKSIMVELVKAIEDKFGVSAAPVVMVSGSAAAAAP---

01aae AQAEEKTEFDVILKSP--GKNKIQVIKVVREIT-GLGLKEAKELVDNAPK---IKEGVP  
04bb --SEEQTEFDVILMSF--GDSKINVIKEVRAIT-GLGLGEAKALVEAAP---KAIEGELS  
05bs --AAEEQSEFDLILAGA--GSQKIKVIKVVREIT-GLGLKEAKELVDNTPKP---LKEGIA  
06cj AAAEEKTEFDIVLTDG--GAKKIEVIKIVRAIT-GLGLKEAKDAVEQTP---STLKEGVA  
07cp PVAAEPTFAVTLLEDVP-ADKKIGVLKVVREIT-GLALKEAKEMTEGLP---KTVKEKTS  
08ct PASAEPTFAVILEDVP-SDKKIGVLKVVREIT-GLALKEAKEMTEGLP---KTVKEKTS  
09der --AAEEKTEFDVILIDA--GASKINVIKEIRGIT-GLGLKEAKDMSEKGG---VLKEGVA  
10ec --AEEKTEFDVILKAA--GANKVAVIKAVRGAT-GLGLKEAKDLVESAP---AALKEGVS  
11hi --AEEKTEFDVILKSA--GANKVAVIKAVRGAT-GLGLKEAKDLVESAP---ANLKEGVS  
12hp AESEEEKTEFNVILADS--GAEKIKVIKVVREIT-GLGLKEAKDATEKTP---HVLKEGVN  
13mg --QEAASEVSVKVTGYA-DNAKLAVLKLYREIT-GVGLMEAKTAVEKLP---CVVKQDIK  
15mp --QEAASEVTVKVTGYT-DNAKLAVLKLYREIA-GVGLMEAKTAVEKLP---CVVKQDIK  
17mtu EAAEEQSEFDVILEAA--GDKKIGVIKVVREIVSGLGLKEAKDLVDGAPK---LLEKVA  
18nm ADAEEKTEFDVVLASA--GDQKVGVIKVVRAIT-GLGLKEAKDIVDGAP---KTIKEGVS  
19pae --AEEQTEFTIIVLAEA--GDKKNVVIKVVREIT-GLGLKEAKAVVDGAP---GVVKEGAS  
21rp EAVTEKTEFEVVLAT--GDKKVEVVIKVDIT-GLGLIEAKLVDEAP---KSIKNVVK  
22syn EAAEEKTEFDVILEEVP-ADKKI AVLKVVRTIT-GLGLKEAKELVESTP---KAIEATG  
23tm AAQEEKTEFDVVLKSF--GQNKI QVIKVVREIT-GLGLKEAKDLVEKAGSPDAVKSQVS  
24tp VAAEEQTEFTVVLKGLAEPGKKI AVIKEVRNIVISGLGLKEAKDLVEGAP---KTLKENVS  
25vib --VEEQTEFNVILAAA--GANKVAVIKAVRGAT-GLGLKEAKALVDGAP---ASVKEAVS  
26xyf --VEEQTEFTVTLKEA--GAKKVEVIKAVRAVT-GLGLKEAKDLTEAG---GILKEAVS

01aae KEEAEQIKKKLEEAGAEVELK-----MLNEFIYPDKIF  
04bb KSDAEELKKKLEAVGAKVEVK-----MMLVEK--FLKDFTIPEKIE  
05bs KEEAEELKAKLEEVGASVEVK-----MIEIEKPKIET  
06cj KAEAEAAKKQLEEAGAKVELK-----MRNI-----TTSAYTPTFT  
07cp KSDAEDTVKKLQDAGAKASFGLMPAKKKAQSVVLGKEKGMSDNHNLLYDKFELPEAVK  
08ct KSDAEDTVKKLQEAQAKAVAKGL-----MSDSSHNLNKKFELPESVK  
09der KDEAEKMAQLEAAGARVELK-----MEQKRPQLKA  
10ec KDDAEALKKALEAGAEVEVK-----MQGS-----VTEFLKPRLVD  
11hi KEEAEALKKELEAGAEVEVK-----MQGS-----VTEFLKPRLVD  
12hp KEEAETIKKKLEEVGAKVEVK-----MKVI-----KTAPLIPSEIK  
13mg PEEAEELKKRFVEVGATVEVK-----MEKF-----LKYEIKVNNNQ  
15mp PEEAEELKKRFVEVGATVEIK-----MEKF-----LKYEIKVNNNEQ  
17mtu KEAADEAKAKLEAAGATVTVK-----MLISQRPRLSE  
18nm KAAEADIQKQLEEAGAKVEIK-----MQNS-----TTEFLKPRQID  
19pae KEEAEAAKALEAGAKVELK-----MQSS-----VNEFLTTPRHID  
21rp KAAEEI KSKLEAAGAKVELK-----MLSLSKN-----WNTLIKPNRVT  
22syn KDDAEA I KQLEEAGKAAVK-----VAQFQIECVES  
23tm KEEAEI KKKLEEAGAEVELK-----MIEFVIPKKLK  
24tp KEEAAKIKESMTAAGALIEIS-----MPRRN-----LLKGFKRPKVLE  
25vib KEEAEALKKELEAGATVEVK-----MQGS-----VTEFLKPRLVD  
26xyf KEEAEKVKLEAAGATVEVK-----MTVT-----VSQVLRPRGPQ

01aae WEEKTDT-----YGRLVVEPLERGFQTTVGNLSRRVLLSSISGTAITAVKIYG-----  
04bb FLKSQDGS-----YGFYIYFPERGFQITIGNTLRRVLLSSIEGYAITAMRVQSNNDSS  
05bs VEISDDAK-----FGKFVVEPLERGFQITLGNLSRRVLLSSIPGAAVTSIQIDG-----  
06cj IENISDT-----VAKISAWPFEIGYITLAHPLRRLLYTSTIGYAPTAIHIDG-----  
07cp MLPVEGLP-----IDKHARFIAEPLERGMGHTLGNALRRALLIGLEAPAIISFAMTG-----  
08ct MSPVEGAVGGIDKVARFVADPLEKGMGHTLGSALRRALLIGLEAPAIISFAMTG-----  
09der --RVDGD-----YGEFVLEPLARGYGVITIGNPIRRILMSSIPGTAVTSVYIED-----  
10ec IEQVSSST-----HAKVTLEPLERGFQHTLGNALRRILLSSMPGCAVTEVEIDG-----  
11hi IEQISST-----HAKVILEPLERGFQHTLGNALRRILLSSMPGCAVTEVEIDG-----  
12hp VLEKEGN-----RVKISLAPFEGYAVTLAHPIRRLLSSVGYAPVGLKIEG-----  
13mg PTNTNPN-----YGIFEVAPLESSGFQITIGNAMRRVLLSCIPGASVFAIASG-----  
15mp -ARANPN-----YGIFEVPLESGFVITIGNAMRRVLLSCIPGASVFAIASG-----  
17mtu DVLTDNR-----SQFVIEPLEPGFGYTLGNLSRRVLLSSIPGAAVTSIRIDG-----  
18nm VNTFSAT-----RAKVMQPFERGFQHTLGNALRRILLSSMNGFAPTEVAIAG-----  
19pae VQVVSQT-----RAKITLEPLERGFQHTLGNALRRILLSSMPGCAVVEAIEDG-----  
21rp YENFPETNN-----KAKIIVEPLERGFGLTLGNAMRRVLLSSIPGAAVTSIKIPA-----  
22syn STRKNQQQ-----YSKFSLEPLDRGQTTVGNALRRVLLSNLPGAAVTAIRIAG-----  
23tm VEEREERD-----YYYSRFSLSPLERGYAITIGNALRRVLLSSIPSLAIVGVRFIKPE-----  
24tp FLSSENSSE-----CYGKFTASPFETGFGTTVGNCLRRVLLSSIPQYAVTGVRITSFDADGV  
25vib IEQISTT-----HAKVTLEPLERGFQHTLGNALRRILLSSMPGCAVTEVEIEG-----  
26xyf IERLTEN-----RAKVLEPLRGYHTLGNALRRVLLSSIPGFAITVEVEIDG-----

01aae ---IYHEFSAIEGVQEDAIELIANLKKIKFLMK-----G-----DSDVEILYQKK--  
04bb SKVVSSEFDLIPGVSEDTLEITIANIKNIHLKLG-----EGEQRKTSISFSVSGK  
05bs ---VLHEFSTIEGVVEDVTIILHIKKLALKIY-----SDEEKTLEIDVQ--  
06cj ---VAHEFDSMRGMLDVALFIINLKKLRFKIK-----G-----DSNKE----IVEFSFK  
07cp ---VLHEYMAIEGVIEDVTNIILNLKGALLKKYPMQ--DSSLGRTTQVLKASISIDASDL  
08ct ---VLHEYMAVEGIIEDVTNIILNLKGSLLKKYPLQ--DCEGGRCSQKLRAISIDASDL  
09der ---VLHEFSTIPGVREDVIRLILNLKELVVKFH-----APGPKTLTLRAQ--  
10ec ---VLHEYSTKEGVQEDILEILLNLKGLAVRVQ-----GKDEVILTNNKS--  
11hi ---VLHEYSSKEGVQEDILEVLLNLKGLAVKQV-----NKDDVILTNNKS--  
12hp ---VHHEFDSLARGVETVSLFIMNLKNIRFIAKALV--GQ--DSSLNQSVVVDYSFK  
13mg ---VKQEFNSVEGVLEDVTEMVLFKQLVVRISDLLFEDGEMIEPPLERVPVKVTAIE--  
15mp ---AKQEFAAVEGMKEDVTEVVLNFKQLVVKISDLLFEDGEMIEPPLERVPLLTVAIE--  
17mtu ---VLHEFTTPGVKEDVTEIILNLKSLVSSSE-----EDEVPTMYLRKQ--  
18nm ---VLHEYSTVDIGQEDVDILLNLKIGIVFKLH-----GRSQVQLVKKS--  
19pae ---VLHEYSAIEGVQEDVIEILLNLKGLAIKLH-----GRDEVTLTLAKK--  
21rp ---IEHEFSSIPGVQEDVSEVILNIKGIEVKMH-----VSEKRI IKLKAM--  
22syn ---VNHEFATILGVREDVLEIMLNKMLVLSKY-----TDQPIGRILTAI--  
23tm ---KYHEYDYIEGVKEDILDILNLKVKQFRIN-----VTVKGTIKMEVEKK--  
24tp AHFISSEFEQIPHVREDTLEILNFKRLRFLLP-----GQAESSTFTYEFR--  
25vib ---VLHEYSTKEGVQEDILEILLNLKGLAVRVAE-----GKDEVITLNNKS--  
26xyf ---VLHEYTTVEGLQEDVLEVLLNLKDVAIRIH-----SGDTATLSLFKQ--

01aae ----GEGEVKASDIKTPPNVEILNPDQYIATITDPNKNELNIEIRVERGRGYVPV---EEM  
 04bb ----DTNVLKASHFERDG-VEVFNKDLVIATLSHD-VNLDLEFQINYGRGYVSS---EQN  
 05bs ----GEGVTAAADITHDSDEILNPDHLIATLGEN-ASFRVRLTAQRGRGYTPA---DAN  
 06cj ----GSKEIYGKDLNNDQ-VEVFNKDAYLATINED-AELKFTLIVEKGIGYVPS---EEI  
 07cp AAANGQKEVTLQDLLQEGDFEAVNPDQVIFTVTQP-IQLEVVLRIAFGRGYTPS---ERI  
 08ct AAAGGQKEVTLGDLQEGDFEAVNPEHVIFTVTRP-MQLEVMLRVAFGRGYSPS---ERI  
 09der ----GEGEVRAFAFEVPTDAEIVNPDLVIANLAED-GKLVMEVRVEEGEGYVSA---DK  
 10ec ----GIGPVTAADITHDGDVEIVKQPHVICHILTDENASISMRIKVRQGRGYVPASTRIHS  
 11hi ----GIGPVVAADITYDGDVEIVNPDHVIICHLTDENASISMRIKVRQGRGYVPASSRTHT  
 12hp ----GPMELRARDLNSEQ-IEIVNPEMPLATINED-AQLNFSLLIYKGMGYVPS---ENT  
 13mg ----KKGAVYAKDLECPAGFEVINKDLYFLSLQKD-MKLTVSVYVKQGRGFTSF---LEN  
 15mp ----KAGPVYAKDLECPAGFEVINKDLYFLSLQTD-KKVTNVVYVKQGRGFVTF---LEN  
 17mtu ----GPCEVTAGDIVPPAGVTVHNPGMHIATLNDK-GKLEVELVVERGRGYVPA---VQN  
 18nm ----GSGVVSAGDIELPHDVEILNPGHVICHLDN-GQIEMEIKVEQGRGYQSVSGRQVV  
 19pae ----GSGVVTAAADIQLDHDEIINGDHVIANLADN-GALNMKLKVARGRGYEPADARQSD  
 21rp ----GPGCVTAGMIDTGHDEIILNPDHVIICHLAKN-KQLEMELTCKVKGKGVLS---TNS  
 22syn ----GPGVTAAQFEVPESEVIDPNQYIATLAEG-AKLEMEFRVERGVGRVVI---ERG  
 23tm ----GPGELVAGDIKTPAGIEVNNPDHLIATLNSK-ADLFFEVYAEVKGKGFVPV---SER  
 24tp ----GAVSLTGKDFAKKFQLEVLSDLLIMEMMDG-AHVEVELHVEFGRGYVPA---ESH  
 25vib ----GSGPVVAGDITHDGDVEIVNPEHVICHILTSDNAAIAMRIKVERGRGYVPASARIHT  
 26xyf ----GAGVTAADIKTDHNVIEIINDGHVICHILTKD-TTINMRLKIERGFGYQPAVRRRP

01aae EA-IGEVGWILVDADFSPVKKVGRFVDNVRVGKKSTYERLTLEIFTNGIKTPDQCMQEA  
 04bb SKYLEEVNVIALDSIFSPIEKVSYSVEDTRVGQRSDYDKLVMEIWTGTVISAKDAIKKAA  
 05bs KRDDQPIGVIPIDSIYTPVSRVSQYQENTRVGGVANYDKLTDVWTDGSTGPKEAIALGS  
 06cj KELINDPKFIALDAFFTPVREATYDIEKVLFDNPDYEVVLTVTDDGQITPNEAFQNAL  
 07cp VLEDKGVYEVILDAAFSPVTLVNYFVEDTRVGQDQDFDLVLIVETDGRVTPKEALAFST  
 08ct VLEERGMEIILDAAFSPVTLVNYFVEDTRVGQDQDFDLVLIVETDGRVTPKEALAFST  
 09der HATKDRINSIPVDAMFSPVRRVAYHVENTRVGQDQDLRLILRVWTDGSGAGQALDRAV  
 10ec EEDERPIGRLLVDACYSPIVERIAYNVEAARVEQRTDLKLVIMETNGTIDPEEAIRRAA  
 11hi QE-ERPIGRLLVDACYSPIVERIAYNVEAARVEQRTDLKLVIMETNGTIDPEEAIRRAA  
 12hp RELMPEG-YMPLDGSFTPIKKVYVEIENVLVEGDPNYEKIIFDIETDGGIDPYKAFLSAV  
 13mg RELINSLGIIATDANFSPVLHCGYEVQVKTSKQKLTDLHFTFKIATNGAIAKADAFAMAA  
 15mp REMINSLGIIATDSNFSPLHCGYEVQELKTSKQKITDHLTFKIIATNGAIAKADAFAMAA  
 17mtu RASGAIEGRIPVDSIYSPVLKVTYKVDATRVQRTDLDKLVLDVETKNSISPRDALASAG  
 18nm RDNRRIGAIQLDASFSPISRVSFEVAPARVEQRTDLKLVLDIETDGSIDPEEAIRRAA  
 19pae EDESRISGRLLQDASFSPVRRVSYYVENARVEQRTDLKLVLDIETNGTIDPEEAIRRAA  
 21rp YEDNLPIGEIAIDALFNPVKSVTYKVENTRVGQVTDYDKLIMFVETNGDVLPEMAVGLAA  
 22syn KDENSLLDFLQIDSVFMPVTKNYNTVEDIRADGMSPKDRLLDIWTNGSIQPREALSEAS  
 23tm EE-RPDVGWIPIDGVFSPVTKVNLFTENVRVGKRTDYDKLILEIWTKKSIRPEALRKA  
 24tp DYADLVGVIPVDAIFSPVLRVRYDIQSCRVGQDGYDQLSLEWTDGTVRPEDAIAEAA  
 25vib EEDERPIGRLLVDATFSPVDKIAYSVEAARVEQRTDLKLVIMETNGTIDPEEAIRRAA  
 26xyf DDENRTIGRLLDASFSPVRRVAYVVEAARVEQRTDLKLVIMETNGTIDAEALRTAA

01aae EILKKHYELLEN-----IFTEKPT----VPQKV  
 04bb SIVREFLFPLVD-----FEDNVNT---SFEKSKS  
 05bs KILTEHLNIFVG-----LTDEAQA--EIMVEKE  
 06cj EAMYKQLSVFDK-----ITNRSVIKNQATSNE  
 07cp QILTKHFSIFEN-----MDEKKIVFEEAISIEKE  
 08ct QILSKHFSVFEK-----MDEKRIVFEEAISVEKE  
 09der EILRDELSVFGNVPEMPALESSYAAATPAAVYDPATATLPASVYDSRPQDLGSLSINPQ  
 10ec TILAEQLEAFVD-----LRDVRQP---EVKEEKP  
 11hi TILAEQLDAFVD-----LRDVRQP---EIKEEKP  
 12hp KVMKQLGVFGF-----RP-IAANTEYSGDYAQRD  
 13mg KILIEHLNPIVS-----VNESIKNL--TIIQEKA  
 15mp KILIEHLNPIVN-----VNESIKAL--NIIQEKA  
 17mtu KTLVELFGLARE-----LNVEAEGI--EIGPSPA  
 18nm RILIDQMSIFAD-----LQGTPVE---EVEEKAP  
 19pae TILQQLAAAFVD-----LKGDSEP---VVEEQED  
 21rp RILQEQLQLFIA-----FEEQEEDK--QVKTDSL  
 22syn DIIANLFIPLKD-----LNELEAAH---SDYQD  
 23tm DILINHFKIVTEGLPE-----LKISEEYIITSEEEAEVPASEHEE  
 24tp KIIKEHFTVFN-----FDETALD----LEDEPE  
 25vib TILAEQLDAFVD-----LRDVRVP---EKEEKP  
 26xyf DILTDQLSVFGD-----FTHDRDG---TVKPASS

01aae AVDELA-EKLSLSIEELDISQRALNSLKRGITIGDLVRMTEDELKSTKNIGRKALAEI  
 04bb ESSN----LLDMSIEKLNLSVRSLNCLAKENVRTLGLISKNAEELSKARNFGKKSLEEI  
 05bs EDQKEK--VLEMTIEELDLSVRSYNCLKRAQINTVQELANKTEEDMMKVRNLGRKSLLEV  
 06cj ENTK----LLQNTIDNLARSYNCLEKAGVYIGELALMSVSELAGLKNLGKKSLEDEI  
 07cp NKDDIL-HKLLVGINEIELSVRSTNCLSNANIETIGELVIMPEPRLLQFRNFGKKSLECEI  
 08ct NKDDIL-HKLLVGINEIELSVRSTNCLSNANIETIGELVIMPEPRLLQFRNFGKKSLECEI  
 09der PFPPTDQ-DTPRVTLGLEGLTTRVLHSLKEEGIDSDVALCALSDDRLKKVPGIGERSLDEI  
 10ec EFPD----ILLRPVDDLELTVRANCLKAEAIHYIGDLVQRTEVELLKTPLNKGKSLTEI  
 11hi EFXP----ILLRPVDDLELTVRANCLKAEAIHYIGDLVQRTEVELLKTPLNKGKSLTEI  
 12hp DAKD----LSAKIESMNLARSFNCLDKIGIKYVGELVLMSEELKGVKNMGKKSLEYDEI  
 13mg EERKVK--SFAKQIEELDFTVRTFNCLKRSIGIHTLQELLSKSLDIREIRNLGKKSEREI  
 15mp EERRVR--SFAKQIEELDFTVRTFNCLKRSIGIHTLQELLSKSLDIREIRNLGKKSEREI  
 17mtu EADHIA--SFALPIDDLTLVRSYNCLKREGVHTVGEVARTESDLDLDIRNFGKKSIDEV  
 18nm PIDP----VLLRPVDDLELTVRANCLKAEIYYIGDLQIRTEVELLKTPLNKGKSLNEI  
 19pae EIDP----ILLRPVDDLELTVRANCLKAEIYYIGDLQIRTEVELLKTPLNKGKSLTEI  
 21rp PFSPP----YLLKRVDELELSVRANCLKNDNIYYIGDLVKRTESDMLRTPNFGKKSLENEI  
 22syn EVNP----ESQPIIEELQLSVRAYNCLKRAQINSVADLLEYSQEDLLEIKNFGKKSLEVE  
 23tm EHRENS--DVYNRKIDLELSVRSLNCLKRAKIIETIGDLSKTEEELLLKINFGKKSLEVE  
 24tp EDDPAVLELNTKIADVDFSVRARNCLTMGIIKTLGELTRISEQTLANTRNVGKKSLEI  
 25vib EFPD----ILLRPVDDLELTVRANCLKAEAIHYIGDL-----  
 26xyf GVDP----VLLRPIDDLTLVRSANCLKAEIYYIGDLIQKTEVELLKTPLNKGKSLTEI

```

01aae KEALHK-LGLELGMNIETQR-----
04bb IEKLGs-YRLYLGMskEDALS---VLSKNVKISE-----
05bs KAKLEE-LGLGLRKDD-----
06cj KNIMES-IGFPVGTskLSDNK--EILKNKIAELKAQNEG-----
07cp KNKLKE-MKLELGMdLTQFGVGLDNVKEKMKWYAEKIR-AKNIKG-----
08ct KNKLKE-MKLELGMdLSQFGVGLDNVKEKMKWYAEKIRSSKNTKG-----
09der KQQLAQ-FGLALRD-----
10ec KDVLAS-RGLSLGMRLenWPP--ASIADe-----
11hi KDVLAS-RGLSLGMRLenWPP--ASIADe-----VLVFLLA
12hp AEKLND-LGYPVGTelsPEQR--ESLKKRLEKLEDKGGND-----
13mg IKKVQE-LGLKFRS-----
15mp IKKVHE-LGLKLRS-----
17mtu KIKLHQ-LGLSLKdSPSPFDp--SEVAGYdVATGTWSTEGAYDEQDYAETEQL-----
18nm KEVLAS-KGLTLGskLEAWPP--VGLEKP-----
19pae KDVLAS-RGLSLGMRLDNWPP--ASLKdDKATA-----
21rp KEILAK-FNLRFGMdVPdWPP--ENIQELSKRYEDSYN-----
22syn IEALQKRLGITLPHEKAKA-----
23tm KEKLKEKfGLELRKGE-----
24tp QGKLQE-YNLRGMADYNHVGVSRLMRQKEEIDEA-----
25vib -----
26xyf KEVLQG-RGLGLGVKLenWPP--PGVSQYGMlg-----

01aae -----MEQEKIRIKLRAyDHRLLDQsVKQIIEtVKRTGGVVGPIPLPRKRKWCVL
04bb -----LIAKDKIRVRLFSFDVKILDQSAESIVKAVQKAKAQIKGPIPLPTKIKKYTVL
05bs -----MAKQKIRIRLKAyDHRILLDQSAEKIVETAKRSGASVSGPIPLPTEKsvYTiL
06cj -----MERIRLKLKAyDHRVLDRTVAaIVEAVKRTGADIRGPIMPTKIKRYTVL
07cp -----MKQKQKQIRIRLKGFDQGLDRSTADIVETAKRTGARVVGPIPLPTKREvYTVL
08ct -----MKQKQKQIRIRLKGFDQGLDQSTANIVETAKRTGARVVGPIPLPTKREvYTVL
09der -----MVAPKIRIKLGFdHKALDQsASKIVdTVRTGADVSGVPVLPTRIRRFTVL
10ec -----MQNQIRIRLKAfDHRLLDQATAIEIVETAKRTGAQVRGPIPLPTRKERFTVL
11hi KFYSSGLMQNQIRIRLKAfDHRLLDQSTAEIVETAKRTGAQVRGPIPLPTRKERFTVL
12hp -----MEKIRLKLKAyDHRVLDRSVVAIVEAVKRSgSEIRGPIPLPTKKNRYTVL
13mg --MNSA--VKYPeLKIKLESYdSTLLDLTIKKIVEVVGvGNIKIGPLPLPTKKEvITII
15mp --MNAANA VKYPeLKIKLESYdSTLLDLTIKKIVEVVGvDVKIKGPIPLPTKKEvITII
17mtu -----VAGQKIRIRLKAyDHEAIDASARKIVETVVRTGASVVGPIPLPTEKNVYCVI
18nm -----MANQKIRIRLKAyDYALIDRSAQEIVETAKRTGAQVRGPIPLPTKIERFNIL
19pae -----MQNQIRIRLKAfDHRLLDQSTQEIIVETAKRTGAQVRGPIPLPTRKERFTVL
21rp -----MKNKIKIRLKSfDHRSLDQATKEIVSAVKRTFATINGPIPLPRKIERFTVN
22syn -----MATLQKQKIRIRLKAfDRRLDTSdKIDVdTNRTNAaAVGPIPLPTKRKIYCVL
23tm -----MPGQKIRIKLKAyDHELDESAKKIVEVAKSTNSKVS GPIPLPTERTLYCVL
24tp -----MARERIRVLCGFdVELVDQSSRAIVHAVQKAGAEVLGPIPLPTRMHKFTVL
25vib -----MQNQIRIRLKAfDYKLIDASTAEIVETAKRTGAQVRGPIPLPTRKERFTVL
26xyf -----MADQKIQIRLKAfDCRLIDRSAGEIVETAKRTGAHVRGPIPLPTKIERCTIL

01aae RSPHKFDQsREHFIEfRSILDIIRfTPQTIEALMEISLPAGVDVEVKMRG-----
04bb RSPHVNKsREQfEMRTHKRLIDIEPTsALMDSLMLKLELPAGVEVDIKQ-----
05bs RAVHYKdSREQfEMRTHKRLIDIVNPTQTVDALMRLDLPSGVDIEIKL-----
06cj KSPHINKdSREQfEIRIHARMLDIVAATPDtVDSLTKLDLAPeVSEVVRAMGK-----
07cp RSPHVdKKSREQfEIRTHKRLVDILDPTGKTIDALKMALPAGVDIKIKAA-----
08ct RSPHVdKKSREQfEIRTHKRLIDILDPTGKTIDALKMALPAGVDIKIKAA-----
09der RSPFKYKdSREHFIEIRTHNRLVDIMNPTKKTIDSLMTLDLPTGVDIEIKTVGGRA-----
10ec ISPHVNkdARDQYEIRTHRLVDIVEPTeKTVdALMRLDLAAGVDVQISLG-----
11hi ISPHVNkdARDQYEIRTHKRLVDIVEPTeKTVdALMRLDLAAGVDVQISLG-----
12hp RSPHVNKdSREQfEIRVYSRLIDII SATPETSdSLMKLDLAPeVDVEVTSMETK-----
13mg RSPHVdKASREQfEKNTHKRLMILVDVNQGGIDSLKKIKIPVGVTIRfSK-----
15mp RSPHVdKASREQfEKNRHKRLMILVDVNQAGIDSLKKIKIPVGVTIRfSK-----
17mtu RSPHYKdSREHFEMRTHKRLIDIDPTPKTVdALMRIDLPAVDVNQ-----MPPAK
18nm RSPHVNKTSREQLEIRTHRLMDIVdWTDKTTDALMKLDLPAGVDVEIKVQ-----
19pae ISPHVNkdARDQYEIRTHKRLVDIVQPTDKTVdALMKLDLAAAGVEVQISLG-----
21rp RSPHVHKSREQfEIRKHKRLVIDDPNPAVDALSKVDLAAAGVDVIELESGE-----
22syn RSPHVdKdSREHFETRTHRRIIDIIQPSsKTTIDALMKLDLPAGVDIEVKL-----
23tm RSPMKHdKdSREHFETRVRHRLIDIIDPSPKTTIDALMRINLPAGVDVEIKL-----
24tp RSPHVNKsREQfEMRTHKRLIDIEPSQEVMMALMGELsAGVDVRKQ-----
25vib ISPHVNkdARDQYEIRTHKRLIDIVEPTDKTVdALMRLDLAAGVDVQISLG-----
26xyf VSPHADkdARDQYETCTYKRVLyIVDPNDKTVdALMKLELAAGVDVQIKLT-----

01aae -----MAKkkkKQRQVTKAIVHIHTfFNNTIVNVTDtQGNtIAWASGGTVGfKGt
04bb LEGKLSAKLSTNSKKKIKRNIeGNNVIQATfFNNTIVTVSDIKGNALAWASAGGMGfKGa
05bs ---MAAARKsNTRKRVRKKNIESGIAHIRSTfFNNTIVTITDTHGNALsWSSAGALGfRGs
06cj ---MAKRKI-VKKKVVKKNIAKGIVYISATfFNNTMVITdEMGNALAWSSAGGLGfKGs
07cp --MVKNQAQAKKsVKRQQLKNIPSGVVHVKAfFNNTIVSITDPAGNVISWASAGKVGSgS
08ct --LVKNQAQ--KKGVKRKQVKNIPSGVVHVKAfFNNTIVTIDPAGNVISWASAGKVGSgS
09der ---MAKPTKGKAPRRSRRNISAGRAYVHASfFNNTIVTIDLDGNSVWASGGTIGYKGS
10ec ---MAKAP--IRARKVRKQVSDGVAHIHASfFNNTIVTIDRQGNALGWATAGSGfRGs
11hi ---MAKTP--VRARKRVKKQVVDGVRHIHASfFNNTIVTIDRQGNALAWATAGSGfRGs
12hp ---MAKRNVtAKKKVVKKNIArGVVYISATfFNNTNITITdEMGNVcWSTAGGLGfKGs
13mg ---MAK-----KKKINVPsGLIHVSCSPNNTIVSATDPsGNVLCWASSGTVGfKGf
15mp ---MAK-----KKKINVSsGIHVSCSPNNTIVSAsDPsGNVLCWASSGTMGfKGs
17mtu KGPATsARKGQKTRRREKKNVPHGAHIKSTfFNNTIVTIDPQGNVIGWASSGHVfKGs
18nm ---MAKANTAsRVKKVRKTVSEGIHVHVHASfFNNTIITIDRQGNALSWATSGGAGfKGs
19pae ---MAKP--AARPRKKVKKTVVDGIAHIHASfFNNTIVTIDRQGNALSWATSGSGfRGs
21rp -----MNQTIKVKKKKKTITLGVVHIRASfFNNTIVTfDIQGNtISSAGGNGfKGa
22syn ---MARPTrKTGPgKAKKNVPSGVAHIQSTfFNNTIVTISDIRGDVSWASAGSSGfKGa
23tm ---MARKGGSsKKQKKSfDYGVVHIKSTfFNNTIITLTDKdGNtLWASGGTVGfEGT
24tp -----VAVTKKRKEKKNVYEGNNVIQATfFNNTIITVTDLQGNALSWASSGLGfNGa
25vib ---MAKQP--TRARKVRKQVADGVAHIHASfFNNTIVTIDRQGNALAWATAGSGfRGs
26xyf ---MAKQS--VVKIKKKVKRVIDTGVAHIHASfFNNTIVTIDRQGNALfWCTSGASGfRGs

```

01aae RKSTPYAAQLAAQKAMKEAK-EHGVQEVEIWKGPAGRE--SAVRAVF-ASGVKVTAIR  
04bb KKSTPYAAQITAESALNKVR-DFGINYHVYIKGPGIGRE--SAIRAIG-SIGMTVKSIS  
05bs RKSTPFAAQMAETAAGKSI-EHGLKTELVTKGPGSGRE--AIRALQ-AAGLEVTAIR  
06cj KKSTPYAAQVADALNKAK-EHGIKEVGIVKQPGSGRE--TAVKSVGMEGKVTFLK  
07cp RKSSAFAATVAAQDAAKTAM-NSGLKEVEVCLKGTGAGRE--SAVRALI-SAGLVSVIR  
08ct RKSSAFAATVAAQDAAKAAM-SSGLKEVEVCLKGTGAGRE--SAVRALI-SSGLIVSVIR  
09der KKSTPYAAQLAADAVKKAQTSFGMAADVIVRSGSGRE--QAIRAIQ-ASGLIVRSIM  
10ec RKSTPFAAQVAERCAVAVK-EYGIKNLEVMVKGPAGRE--STIRALN-AAGFRITNIT  
11hi RKSTPFAAQVAERCAEIVK-EFGLKNLEVMVKGPAGRE--STIRALN-AAGFRITNIT  
12hp KKSTPYAAQVAVESALS KAK-EHGVKEVGIVKQPGSGRE--TAIKSVGATEGIKVLWIK  
13mg RKSTPYAAGVAADKVAKTAK-EMGMGSKVMYKGTGRGKD--TTIRSA-NAGITITEIN  
15mp RKSTPYAAGIAADKVAKTAK-EMGMATVKLFVKGTRGKD--TAIRSA-NAGLSITEIN  
17mtu RKSTPFAAQLAENPARKAQ-DHGVKRVDFVVKGPAGRENPDVGCAGR-RS-WRWARSR  
18nm RKSTPFAAQVAEAAQKVAQ-EYGVKNLEVRIGKGPAGRE--SSVRALN-ALGFKITSIT  
19pae RKSTPFAAQVAERAGQAL-EYGLKNLDVNVKGPAGRE--SAVRALN-ACGYKIASIT  
21rp RKATPYAAQVTIDKASEKAK-ECGLKTSIRIRIGPAGRE--SAMRALF-GQNFVVTISIL  
22syn KKSTPYAAQTAADSARRAM-EQGMRLQLEVMVSGPGAGRE--TAIRALQ-GAGLEITLIR  
23tm RKSTPYAAQLAADKVAAREAL-RMGIKKVDVLFVKGPGAGRE--PAIRTLQ-GAGLEINQIK  
24tp KKSTPFAAQTVAAEAQKVAQ-QCGLREVHVVFVKGPAGRE--SAIRMLG-TMLRVRSIR  
25vib RKSTPFAAQVAERCAEMAK-EYGLKNLEVMVKGPAGRE--STVRALN-AAGFRITNIV  
26xyf RKCTPFAAQVAEAKGRAVL-DYGMKSLEVRINGPAGRE--SAVRSLN-NVGYKITNII

01aae DVTPIPHNCRPP-ARRRV-----MPTFNQLVKYGREKKKSKAPALQ  
04bb DITPIPHNGCRPK-KTRRV-----MPTINQLIRKPKSQTEKTASPALQ  
05bs DVTPVPHNGCRPP-KRRRV-----MPTINQLIRKGRVSKVENSKSPALN  
06cj DITPLAHNGCRPP-KRRRV-----MPTINQLVRKERRKVKLEKSKSPALK  
07cp DETPVPHNCRPP-KRRRV-----MPTINQLIRKRRKSSLARKKSPALQ  
08ct DETPVPHNCRPP-KRRRV-----MPTINQLIRKRRKSGATRKSPALQ  
09der DDSPVPHNGCRPK-KKFRA-----MPTTQQLLRKGRKVLQKSKVPALK  
10ec DVTPIPHNCRPP-KKRRV-----MATVNLVRKPRARKVAKSNVPALE  
11hi DVTPIPHNCRPP-KKRRV-----MATINQLVRKPRVKVVKSNVPALE  
12hp DITPLPHNGCRPP-KRRRV-----VPTINQLIRKERRKVKKTKSPALV  
13mg EKTPIPHNGCKLLSVRANQNNLWKNF-----MATIAQLIRKPRKVKVSKSPALH  
15mp EKTPIPHNGCKPP-KRPR-----MATIAQLIRKPRKVKVSKSPALH  
17mtu MSPPTRIITVSGPP-NHRRV-----MPTIQLVRKGRDRKISKVKTAALK  
18nm DVTPPLPHNGCRPP-KKRRV-----MPTINQLVRKGRQKPVYVNVKPALE  
19pae DVTPIPHNCRPP-KKRRV-----MATINQLVRKPRKRMVDKSDVPALQ  
21rp DVSSIAHNVRPP-KRRRV-----MPTYNLVRFRGRSKTKTKSPALE  
22syn DVTPIPHNCRPP-KRRRV-----MPTIQLIRKRSKVKKTKSPALK  
23tm DVTPIPFNCRPP-KRRRV-----MKMPTINQLIRYGRPKKKKSKAPALQ  
24tp DITPIPHNGCRPP-KTRRI-----MPTINQLTRIGRAVFSRTKSPALQ  
25vib DATPIPHNGCRPP-KKRRV-----MATINQLVRKPRKQVVKSNVPALA  
26xyf DVTPIPHNCRPP-KKRRV-----MLYLEESTMATINQLVRKPRQASTVKSASPALD

01aae G-----CPQKRGVCRVYVTPPKKPNALRKVARVRLSNGIEVTAYIPGEGH  
04bb N-----CPQRRGICTRVMTVTPKKPNALRKVARVRLSNGIEVTAYIPGIGH  
05bs KGYNSFFKKEHTNVSSPQKRGVCTRVGTMTPKKPNALRKVARVRLTNGIEVTAYIPGIGH  
06cj N-----CPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
07cp K-----CPQKRGVCLQVKTTPKKPNALRKVARVRLSNGIEVTAYIPGEGH  
08ct K-----SPQKRGVCLQVKTTPKKPNALRKVARVRLSNGIEVTAYIPGEGH  
09der G-----SPFRRGVCTRVYTTTPKKPNALRKVARVRLSNGIEVTAYIPGEGH  
10ec A-----CPQKRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
11hi A-----CPQKRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
12hp E-----CPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
13mg YNLNLLNKTTNVYSPKRGVCTRVGTMTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
15mp YNLNLLNKVTNVYSPKRGVCTRVGTMTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
17mtu G-----SPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
18nm A-----CPQKRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
19pae N-----CPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
21rp S-----NPFKSGVCLVKTTPPKPNALRKVARVRLTNGIEVTAYIPGEGH  
22syn Q-----CPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGIGH  
23tm G-----NPQKRGVCIKVTMTPKKPNALRKVARVRLSNGIEVTAYIPGIGH  
24tp A-----CPQKRGVCTRVMTVTPPKKPNALRKVARVRLSNGIEVTAYIPGIGH  
25vib A-----CPQKRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH  
26xyf K-----CPQRRGVCTRVYTTTPKKPNALRKVARVRLTNGIEVTAYIPGEGH

01aae NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAAGVEGRQRSKYGTGKRPKEEGG----  
04bb NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
05bs NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
06cj NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
07cp NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
08ct NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
09der NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
10ec NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
11hi NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
12hp NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
13mg NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
15mp NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
17mtu NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
18nm NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
19pae NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
21rp SVKEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
22syn NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
23tm NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
24tp NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
25vib NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----  
26xyf NLQEHSLVLRGGRVKDLPGVRYKIIRGALDAGVNNRKRGRSKYGTGKRPKPA-----



01aae ---MARIAGVDLPNNKRLVALTYIYGIGWSRAREICEKTGIPCTKRVGELTPDELNTRL  
 04bb ---MARISGIDLNNKQLKIALTSYIGIGRTRALEVCNKSSISPSKIAKLDNDENVNRLR  
 05bs ---MARIAGVDIPRDKRVVISLTYIFGIGRTTAQQVLKEAGVSEDTRVRDLTEELGKIR  
 06cj ---MARIAGVDLPKKRIEYGLTYIYGIGLFTSRKILDKVGISYDKRVHSELSEDEAAAIR  
 07cp ---MPRIIGIDIPAKKKLISLTYIYGIGSARSDEIIKKLKDPEARASELTEEEVGRLN  
 08ct ---MPRIIGIDIPAKKKLISLTYIYGIGLPAKSEIARLQNPPEARAAELTEEEVGRLN  
 09der ---MARIAGIDLPREKRVEIALTYIYGIGLTSRKEILARTGVSPDTRVKNLSEAEQSTLR  
 10ec ---VARIAGINIPDHKHAVIALTSYIGVGKTRSKAILAAAGIAEDVKISELSEGGIDTLR  
 11hi ---VARIAGINIPDHKHAVIALTYIYGIGKTRSAICAAAGIAEDVKIRELSEEGIDKLR  
 12hp KKKMARIAGVDLPKKRVEYALTYIYGIGLSSREILEAVGISFDKRVHSELSEDEVSSIA  
 13mg ---MARILGIDIPNQKRIEIALTYIYGIGLSSAKTILKKAKINPDKRVKDLSEELVAIR  
 15mp ---MARILGIDIPNQKRIEIALTYIYGIGLSSQAILKQANINPDKRVKDLTEEEFVAIR  
 17mtu ---MDRLVGVDLPKRHMEVALTYIYGIGRTRSNEILAAATGIDRDLTRDLTEELIHLR  
 18nm ---MARIAGVNIPNNAHIVIGLQAIYIGIGATRAKLICEAANIAPDTKAKDLDETQLDALR  
 19pae ---MARIAGVNIPDNKHTVISLTYIYGIGVGRTTAQSCAATGVNPAKIKDLSDEQIDQLR  
 21rp ---VARIASVNIPDNKRLVVSLEYIGLSTMAAEICNKAISKDKKVKVLTQELISLR  
 22syn ---VARIAGVDLPDKRVEIALTYIYGIGLSSHEILDATGVSPDVRVLDLSEEDALKLR  
 23tm ---MARIAGVDLPNNKVVVALTYIYGIGRSRSEILKNTGIDPEKRVGDLTDEEISKIT  
 24tp ---MARIAGVDLPN-KHVSVALTYIYGIGRSRARTICEKARISSACLINDLSQDELAVVR  
 25vib ---MARIAGINIPDHKHAVIALTYIYGIGKTRSKAILADLGAESVKISELTEEGIDQLR  
 26xyf ---MARIAGVNLAIQKHVWIGLSYIGIGRTRSCKVCAANVAITYTKIRDLSEPEIERLR

01aae KFIDENYKVEGDLRREVQLNKKLMDMGCYRGIRHARGLPVRGQQTRTNARTRGKRRKT  
 04bb KVIESDYIVEGDLRREVAMSIKRLMDIACYRGVRHRKGLPLRGQRTKTNARTRGKRRKT  
 05bs DIIDK-LKVEGDLRREVSLNIKRLIEICTYRGIRHRRGLPVRGQNSKNNARTRGKPRRTV  
 06cj KEIQENYMEGDLRKKVAMDIKALMDLGSFRLRHRKGLPVRGQTKTNARTRGKRRKT  
 07cp SLQSEYTVEGDLRREVQSDIKRLIAHISYRGQRHRLSLPVRGQRTKTNSRTRKGKRRKT  
 08ct ALLQSDYVVEGDLRREVQSDIKRLITTHAYRGQRHRLSLPVRGQRTKTNSRTRKGKRRKT  
 09der EAIKTYKVEGDLRREVQSNIKRLMDIGAYRGLRHRKGLPVRGQRTKTNARTRGKPKKT  
 10ec DEVAK-FVVEGDLRREISMSIKRLMDLGCYRGLRHRKGLPVRGQRTKTNARTRGKPKPI  
 11hi DEVGK-FTVEGDLRREVTLNIKRLDLGCYRGLRHRKGLPVRGQRTKTNARTRGKRVHSRS  
 12hp KKIQQSYLVEGDLRKKVQMDIKSLMDLGNRYGIRHRKGLPVRGQTKNNARTRGKPKKT  
 13mg NAASG-YKIEGDLRREIALNLIKHLTEIGSWKGIHRKKNLPVRGQRTTNARTRGKPKKT  
 15mp NVASA-YKIEGDLRREIALNLIKHLSEIGAWRGLRHRKKNLPVRGQRTTNARTRGKPKKT  
 17mtu DYIEANLVEGDLRREVQADIRKIEIGCYQGLRHRGMPVRGQRTKTNARTRGKPKAHH  
 18nm DQVAK-YEVEGDLRREVMSIKRLMDMGCYRGFRHRRGLPCRGGQRTTNARTRGKPKAI  
 19pae NEVAK-ITTEGDLRREINMNIKRLMDLGCYRGLRHRKGLPVRGQRTKTNARTRGKPKPI  
 21rp NIIENYKVEGDLRREVTLNKKKKDIRCYQGLRHRKGLPVRGQNTHSNARTRGKKAIAI  
 22syn TYIDENYIEGDLRWEAMNIKRLGDIGYRGRHRKGLPVRGQRTTNARTRGKRRRLT  
 23tm KYIQDHFKEGDLRSEVERNIRRLIEIGCYRGIRHKLGLPVRGQRTTNARTRGKPRPSR  
 24tp AIIDREYKVEGDLRTEVALNIKRLMDIGCYRGLRHRKGLPVRGQRTTNARTRGKRRKT  
 25vib DGVAK-YTVEGDLRREVSMNIKRLMDLGCYRGLRHRKGLPVRGQRTTNARTRGKPKPI  
 26xyf VEVGK-VYIEGDLRREVGMATIKRLMDLNCYRGLRHRKGLPVRGQRTTNARTRGKPKAI

01aae GGTKKAKAK-----MTEQKQ  
 04bb ANKKIASK-----  
 05bs ANKKK-----  
 06cj GAKS-----  
 07cp AGKKK-----M  
 08ct AGKKK-----M  
 09der AGKKKATRK-----  
 10ec KK-----  
 11hi KNSRGK-----  
 12hp GSK-----  
 13mg ANKKIESK-----  
 15mp ANKKIESK-----  
 17mtu RRQEEG---VMAEAKTGAKAAPRVAKAAKAAAPKKAAPDAEAIGAANAANVKGPKHPTPR  
 18nm AGKK-----  
 19pae RK-----  
 21rp AGKKKTVK-----  
 22syn AGKKKTPAKK-----  
 23tm IKTKKSS-----  
 24tp AGKKK-----  
 25vib KK-----  
 26xyf KK-----MMN

01aae QKKWHEKRKHLGVVVSDKMDKTIVVKVDRKVPHPYIGKHIISKKKYHAHDE-HNECRVG  
 04bb --MARENKKEIGKVVSDKMSKTIVVEIVQRKMHPYIHKYLVKSKKYKAHDE-KEVSKVG  
 05bs -MSERNQRKVYQGRVSDKMDKTITVVETYKKHHTLYGKRVKYSKKFKAHDE-NNQAKIG  
 06cj ---MAFKREIQGVVVKIAGEKTASVLVERKVHPYRKIVKRFKKYLIHDE-RNEVKVG  
 07cp ASEPRGSRKVIGVVVSAMKMEKTIVVRVERIFSHPQYLVKVRSSKKYYAHT--LKVSEG  
 08ct ASDVRGRRKTKIGVVVSKMEKTIVVRVERVYSHQYAKVVRDSSKKYYAHNE--LDVKEG  
 09der -----MKTFTFGVVVSDKADKTIVSVKVERFAHPLYGKVVTRSHKYYAAHDE--NNEYKIG  
 10ec ---MTDKIRTLQGRVSDKMEKSIIVVAIERFVKHPYIGKFIKRTTKLHVHDE--NNECGIG  
 11hi ---MTDKIRSVQGVSDKMEKSFVVAIERVKHPLYGKFIIRRTTKLHVHDE--NNEAKVG  
 12hp MNTKEPHKRLVQGVISKFAEKSAVILVERKVVEKYRKIVKKFKKYTIHDE--NNQKVG  
 13mg --MKRNQRKQLIGTVVSTKNAKTATVKVTSRFKHPLYHKSIVIRHKKYHVHNGFELVANDG  
 15mp --MKRNQRKVLIGIVKSTKNAKTATVQVESRFKHPLYHKSIVVRHKKYQAHNEGEVLAKDG  
 17mtu TPKPRGRKTRIGVVVSDKMDKTIVVELEDMRHPLYGKIRTTTKVKYKAHDE-DSVAGIG  
 18nm -MSETKNVRTLQGVVSDKMDKTIVTLVERKVHPLYGKIRLSTKIHADHE--NNQYKIG  
 19pae MAEAKQTVRTLGRVSDKMDKTIVTLERRVKHPYIGKYVKRSTKLHAHDE-SNQCRIG  
 21rp -----MPKRVLQGVVSSKTDKTIVTKVERKFKHPYIKKFKVKSKKYAAHDI-ENKYKEG  
 22syn -----MAIKERVGIVVSNKMDKTIVVAVESRSPHPYKGIIVVTKKFKFAHDE-ENQKQEG  
 23tm -----MPKRLTGIVVSDKMDKTIVVAEVLQHPYKYYKRTTKKYKAHDE-RNECKIG  
 24tp --VKRPERRTLVGLVSDKMDKTIVTRITTKKLHALYKKYVSRSSKKYQAHDE-ENTARAG  
 25vib ---MSDKIRTLQGRVSDKMDKSIIVVAIERVMKHPYIGKFKVKTITKVHAHDE--NNECGIG  
 26xyf DHNERKPLRTIKGRVSNKMDKTIVTLVERQIKHALYKGYIKRSTKLHAHDA--DDLCEG

```

01aae DIVMIRETRPLSKTRWVVVKILQRRARPPEEEIQKQEGGEQ-----VREAIK
04bb DKVKIIEVRPISKDKRWSLVEVLEKLK-----
05bs DIVKIMETRPLSATKRFRLEVVEEAVII-----
06cj DTVVAVECRPLSKRKSFRKLSVLTATGE-----
07cp DKVKIQETRPLSKLKRWRVIEHVGVS-----
08ct DTVRIQETRPLSKTRWRVVGVRN-----
09der DRVEIIAVRPISKTKTWKTKLIERPRGIETTLAETEVAGGEA-----
10ec DVVEIRECRPLSKTKSWTLVRVVEKAVL-----
11hi DTVEIRECRPLSKTKSWTLVRVVEKAVIA-----
12hp DFVSAIECRPLSKTKSFTLKEILVVG-----
13mg DRVQIIETRPLSALKRWRIVKIIERAK-----
15mp DKVQIVETRPLSATKRFRIAKIIERAK-----
17mtu DRVSLMETRPLSATKRWRVLEILEKAK-----
18nm DVVVISERPLSKTKSWVSELVEKARSI-----
19pae DLVTIRETRPLAKTKAWTLVDIVERAVEV-----
21rp DKVSIVESRPISKTKTWVVLNLE-----
22syn DKVRIQETRPLSKTRWQVINIMSHSS-----
23tm DVVEIEETRPLSKAKRWRVRIIQRFE-PERVVKEKEDIQEEIEAVEGKGGVES-----
24tp DVVRIAESRPLSRKRWRVLEIVEERAK-----
25vib DTVEIRECRPLSKTKSWTLVKIVEKAKM-----
26xyf DVLMTVEVAPISKTKNRVRVEIVARSD-----

01aae RYGSFELLKGKNTLPELEALLEERKLVLENLKKQLKEAHKGPKEAEGDEKLKELIREV
04bb -----
05bs -----
06cj -----
07cp -----
08ct -----
09der -----
10ec -----
11hi -----
12hp -----
13mg -----
15mp -----
17mtu -----
18nm -----
19pae -----
21rp -----
22syn -----
23tm -----
24tp -----
25vib -----
26xyf -----

01aae NKAQAEVRALEIIVNRVRKYEELEYAQYQMTEKKAYVDPKLVWRIRKMNETGERK---VV
04bb -----VARSIKKGPPIEKSLYQKVLSS--FGSEK-RVVI
05bs -----MARSLKKGPFVDGHLMTKIEKLNEDKKQ---VV
06cj -----MARSLKKGPFVDHVMKKVIAAKANDNK---PI
07cp -----MSRSLRKGPFVDHLLKKVRAMNIEKKT---PI
08ct -----MSRSLRKGPFVDHLLKKVRDMNALEKKT---PI
09der -----MPRSLKKGPFVEDHLLKKVDAQNDKKDKR---VI
10ec -----MPRSLKKGPFIDLHLLKKVEKAVESGDKK---PL
11hi -----MPRSLKKGPFIDLHLLKKVEKAVESGDKK---PI
12hp -----MSRSIKKGPFIIDHLLMKKTLKAKEGKDNR---PI
13mg -----MSRSSKKGAFFVDAHLLKKVIEMNKQAKKK---PI
15mp -----MSRSKKGAFFVDAHLLKKVIDMKNQEKKK---PI
17mtu -----MPRSLKKGPFVDEHLLKKVDVQNEKNTKQ---VI
18nm -----MARSLKKGPYVDLHLLKKVDAARANDKR---PI
19pae -----MPRSLKKGPFIDLHLLKKVEKAVESKDRK---PI
21rp -----MARSIWKGPFVDGYLIKVKQLMKSGKSE---MI
22syn -----MGRSLKKGPFVAASLRKIDKLNKGDKQ---VV
23tm -----MGRSRKKGPYVDRKLEIKRKLNETGEKK---VI
24tp -----MSRSVKKGPFVDKKLYKRVVEMNKAANQRNKKVI
25vib -----MPRSLKKGPFIDLHLLKKVEKAVESGDKK---PI
26xyf -----MPRSTKKGPFVDHLLIKKVESA--AGSKR---PI

01aae RTYSRATTIIPFVGHTIAVHNGKTFVPVYITQDMVGHLGEFAPTRTFKGHPEKTAKVV
04bb KTYSRSTIIPFVSLTISVYNGKTFIPIYITEDLVGHKLGEFSPTFRGHAK--SDKK
05bs KTWRRSTIIPQFIGHTIAVYDGRKHVPVFISEDMVGHLGEFAPTRTYKGHSHD--DKK
06cj KTWRRSTIIPDMIGLTFVHNGKSFIPVYITENHIGYKLGEFAPTRTFKGHGK-SVQKK
07cp KTWRRSMITPEMIGHTFEVHNGKFLTFVSETMVGHLGEFSPTRFKSHS---VKKG
08ct KTWRRSMITPEMIGHTFEVHNGKFLTFVSETMVGHLGEFSPTRFKSHS---VKKG
09der KTWRRSTIIPDMIGHTIAVYNGKQHVFPVFVSEMQMVGHLGEFSPTRFYGHGA---DKN
10ec RTWRRSTIIPNMIGLTIHVHNGRQHVFPVFTDMVGHLGEFAPTRTYRGHAA---DKK
11hi KTWRRSMIIPMIGLTIHVHNGRQHVFPVVDSEMQMVGHLGEFAPTRTYRGHAA---DKK
12hp KTWRRSTIIPDMIGTFYVHNGRVFPVYITENHVGKLGEFAPTRTFKGHGK-SVQKK
13mg KTWRRSTIIPFVGNTFVHNGKTFINNVYITDDMVGHKLGEFSPTRFKQHT---ANR-
15mp KTWRRSTIIPFVGNTFAVHNGKTFINNVYITDDMVGHKLGEFSPTRFKQHT---ANR-
17mtu KTWRRSTIIPDFIGHTFAVHNGKQHVFPVFTSMVGHLGEFAPTRTFKGHIKD---DRK
18nm KTWRRSTIIPDFIGLTIHVHNGRTHVPVFISDMVGHLGEFSLTRTFKGHLA---DKK
19pae KTWRRSMILPHMGLTIHVHNGRQHVFPVVDSEMQMVGHLGEFAATRTYRGHAA---DKK
21rp KTWRRSTIIPFVGTFVHNGKTFIPVYITENHVGKLGEFAPTRTFKGHG--ADKK
22syn KTWRRSTIIPQMGVHTIAVHNGRQHVFPVSEMQMVGHLGEFAPTRTFRSHSKS---DKK
23tm KTWRRSMIIPKMGVHTIAVYNGKHIPVYITENMIGHRLGEFAPTRFRGHADKAKKG
24tp KYSYRSTIIPDMVGHTISVHNGKSWIPVYITEFVGHKLGEFSPTRFVGHSG--SDKK
25vib KTWRRSMIIPMIGLTIHVHNGRQHVFPVFTDMVGHLGEFAPTRTYRGHAA---DKK
26xyf KTCRRSVILPQMGHTIAVHNGKNYPVVINENMVGHKLGEFSITRVFGHG---GDKK

```



01aae KKK---MGQKTHPIGFRLG--VIKEWPSKWYAPK-KEYSKLLHEDLKIKNYIKERYK--  
 04bb -GRK---MGQKVHPYSLRVK--INKDWKSKWYFDK-KLYSAILHEDFLIRREIMKFLKGI  
 05bs TRR----VQQKYNPVGLRIG--VIRDWESKWYAG--KDYADFLHEDLKIREYISKRLS--  
 06cj IGK---MGQKVNPIGLRLG--INRNWESRWFPFK--ANLVENIGEDYKIRAFLLKKLY--  
 07cp -----MGQKGCPIGFRTG--VTKKWRSWLYGNN-QEFGKFLIEDVIRQFLKKKPS--  
 08ct -----MGQKGCPIGFRTA--VTKKWRSWLYGNN-QEFGKFLIEDVIREFLKKKPS--  
 09der AKGSKKKMGKNINPNPGRFG--VTKGWNSRWYAGK-KQYASLLKEDEQIRQLINKKLA--  
 10ec AKKK---MGQKVHPNGIRLG--IVKPWNSTWFANT-KEFADNLDSDFKVRQYLTKELA--  
 11hi AKK---MGQKVHPNGIRLG--IVKPWNSTWFANT-QDFADNLEGDFKVRKFLNKELA--  
 12hp IGK---MGQKYNPVGLRLG--INRNWTSRWFPFA-RTAPSNIDEDNKIRKFLKKELY--  
 13mg -----MGQKVNSNGLRFG--INKNWSIRWTASSNQQTATWLVQDEKIRNLFYINRY--  
 15mp -----MGQKVNSNGLRFG--INKNWSIRWTANSHAQTAKWILDEKIRNLFYVNYR--  
 17mtu SKRR---VQQKINPHGFRLG--ITTDWKSRYAD--KQYAEYVKEDVAIRLLSSGLE--  
 18nm AKKK---MGQKINPTGFRLA--VTKDWASKWFAKS-TDFSTVLKQDIDVRNLYRQKLA--  
 19pae AKR---MGQKVHPNGIRLG--IVKEHTSVWYADR-KNYADYLFADLKVREYLQDKLK--  
 21rp VKRK---MGQKVCAGHFRVGPLTKDWDKSYLAE--KHYKTLFIQDLKIRDLINKWFN--  
 22syn ARK---MGQKIHPVGFRLG--ITDKHKSRYADP-KRYPELLQEDHKIRQYIEKTLN--  
 23tm ELKK---VQQKVHPGFRLG--LSADWQAKWFNE--KNYKEWLEDEEIRKIINKKY--  
 24tp VGR---MGQKVSPIGLRLG--INKVSSRWYAGP-REYALHLEDLRISMIRSFPECK  
 25vib AKKK---MGQKVHPNGIRLG--IVKPWNATWFANT-KDFADNLDGDFKVRQYLSKELA--  
 26xyf SGK---MGHKVHPGIRLG--ISADWNSKWYANK-AEFGYLAADLKVRQVLRKKMS--  
  
 01aae VAGVSKVEIER-IVDKVRVKIHTARPAIVIGRRGQEVENLKKTKLEMLPG--KEITISVV  
 04bb KFDISDIEIIRNPNQKVTVVIVTPRPGSVIGLKGNSLEKIQGLTKKIS---KKISIKIK  
 05bs DASVSKVEIER-AANRVNITIHTAKPGMVIGKGGSEVALRKALNSLT--GKRVINIL  
 06cj YAGISQILVER-TAKKLRVTVAARPGLIIGKKGSDVDNLRKELQDLIG---KDVNINIK  
 07cp CQGAAGFVVR--MSGKIEVTIQTARPLVIGKGAEDVLLKEELRALTG---KEVWLEIA  
 08ct CQGAAGFVVR--MSGKIEVTIHTARPLVIGKGAESVLLKAEKLLTD---KDVWVEIA  
 09der AAGIARIEIER-AGQQVNVIIISAAKPGVVIKGGSEIKELRGDIERLV--SAGTVAVNVA  
 10ec KASVSRIVIER-PAKSIRVTIHTARPGVIGKKGEDVEKLRKVADIA--GVP AQINIA  
 11hi SASVSRITIER-PAKSIRVTIHTARPGVIGKKGEDVEKLRNAVSKIA--GVP AQINIA  
 12hp YAGVSEIVIER-AAKLRVTVAARPGLIIGKKGVDIEKVDGLKTLIK---KEVSINIK  
 13mg NAQVSNVEIER-TQTTVDVYVYAAQPAALLIGSENKNIKITKMIQIIVG-RKIKLDTIN  
 15mp NAQVSNVEIER-TQATVDVYVYAAQPAFLLIGSENKNIKITKIQIIG-RTNLDLTIN  
 17mtu RAGIADVEIERTRDR-VRVDIHTARPGVIGRRGTEADRIADLEKLTG---KQVQLNIL  
 18nm NASVGRVIER-PAKSARITIHSARPGVVIKKGEDIEVLKRDQLVLM---GVPVHNIE  
 19pae SASVSRIDIER-PAQTARITIHTARPGVIGKKGEDVEKLRQDLTKQM---GVPVHNIE  
 21rp QAQISRVLIERPSNKSIIINAKKNPNIIGKNGTEIDKLKKAIENTMTFL--KEYVINIH  
 22syn NAGISDIRIER-KAEQIELGIHTARPGVVGSGGSGIEQLREGQLKLLG-SARQIRVNI  
 23tm HAGISEIYVERPDAERINITVKTARPGIIIGRKGSSEITSLREELERKFN---RRVVINIE  
 24tp NADIAEVEIVR-HPQRVTVMHTARPGVVIKAGVNIKIEKIGAEVQKRLN---KVKQIKVK  
 25vib NASLSRIVIER-PAKSIRVTIHTARPGVVIKKGEDVEKLRAAVAKIA--GVP AQINIA  
 26xyf QAGISKILIER-PSNTACVSMHVARPGVVIKRGEDIEMLRKQVSDIM---GVPVHNIV  
  
 01aae EVRVPDLDAQLVAQDIATQIERRVSHRRAMKRAIDNALKA-GAGKGVQVQKGRIGGAELA  
 04bb EVKRPELDAQIIANGIAKQVENRASYRKLKSSLSTMLK-GAQLKIKIAGRLGGAELA  
 05bs EIKRADLDAQLVADNIAQLENRVSFRAQKQIQRTMRA-GAQGVKTMVSGRLGGAELA  
 06cj EERKAGASQAQLAESVATQLEKRIAFRRAMKKVIQGAQKA-GAGKIKVSVSGRLGGAELA  
 07cp EIKRPELNAKLVAQDIATQIERRVSHRRAMKRAIDNALKA-GAGKGVQVQKGRIGGAELA  
 08ct EVKRPELNAQLVADNIAKQIERRVSHRRAMKRAIDNALKA-GALGVKQVQSGRLGGAELA  
 09der EIPNPNISAPLVALRIAEQIERRAFRRAMKQAAQRVMES-GARGVKVLSGRLGGAELA  
 10ec EVKRPELDAKLVAQDIATQIERRVSHRRAMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 11hi EVKKPELDAKLVAQDIATQIERRVSHRRAMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 12hp EVKHPQADQAQLAESVATQLEKRIAFRRAMKKVMQAALKS-GAGKIKVSVSGRLGGAELA  
 13mg EIGSPMLSSRIIARDIANAIENRVPLRSAMRQALTKVLKA-GANGIKVLSGRLGGAELA  
 15mp EIGSPMLSARIIARDIANAIENRVPLRTAMRQSLIKVLKA-GANGIKVLSGRLGGAELA  
 17mtu EVKNPESQAQLVAQGVAEQLSNRVAFRRAMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 18nm EIRRPDLAQIIADGIAQLEKRVQFRAMKRAMQAMRS-GAGKIKVSVSGRLGGAELA  
 19pae EIRKPELDAMLVAQGVAEQLERRVSHRRAMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 21rp EVRKFNIDAAIVAQTIAAQLEKRVSFRRAMKTAIQASFKQ-GGQGIKRVSGRLGGAELA  
 22syn EVPNADADAALMAEYIQQLEKRVSFRRVVRQALQRAERA-EVKGKIKVSVSGRLGGAELA  
 23tm EIKTPDLDAQLVAESIASRIEKASYKVMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 24tp EIKRMELNAYLVAQNVARQLTARVSFRKCLRQACAGTMKS-GAQGVKIRVSGRLGGAELA  
 25vib EVKRPELDGQLVADSIAQLEKRVSHRRAMKRAIDNALKA-GAGKIKVSVSGRLGGAELA  
 26xyf EVKRPELDAQLVAESVAQLEKRRIMFRAMKRSVGNAMRL-GALGKIKVSVAGRLGGAELA  
  
 01aae RKEWFLVGRMPLQTLRADIDYGFATAYTKYIGILSVKVMYIKGVDVNLGR-----  
 04bb RSFEVKEGRVPLHTLRANIDYGFSEAQTTYGIGVKVWIFKGEVLGRQTNDSAGQVINKK  
 05bs RPEYYSEGTVPLHTLRADIDYATSEADTTYGKLGKVVWYIRGEVLPTRKKNEEGGK---  
 06cj RTEWYLEGRVPLHTLRADIDYGFAEARTTYGIGVKVWIFKGEVLHGGMQPEKTEESAPA  
 07cp RSEWYKNGRVPLHTLRADIDYATCAEATTYGIGIKVWINLGENSESTTPNNAPASAAA  
 08ct RSEWYKNGRVPLHTLRADIDYATCAEATTYGIGIKVWINLGEMKAVPAANHAGAATAA  
 09der RTEKYLEGRVPLHTLRADIDYGTALARTTYGIGIKVWIFVNGEVIQKTEFTARPQRDR  
 10ec RTEWYREGRVPLHTLRADIDYNTSEAHTTYGIGVKVWIFKGEILGGMAAVEQPEKPAAG  
 11hi RSEWYREGRVPLHTLRADIDYNTSEAHTTYGIGVKVWIFKGEILGGMAAVEQPEKPAD  
 12hp RTEWYMEGRVPLHTLRADIDYGFAEAMTYGIGVKVWIFKGEVLQKGIQFEKKEEAKEE  
 13mg RDKMYIEGNMPLSTLRADIDYAFKAKTTYGIGVKVWINRGMIMYAKGLNRTPAHILHPQ  
 15mp RDKMYIEGNMPLSTLRADIDYALEKAQTTYGIGVKVWINRGMIMYAKGLNRTPAHILHPQ  
 17mtu RSEFYREGRVPLHTLRADIDYGLYEAQTTGIRGKVVWYIKGDIQVGGKRELAAPAGAD  
 18nm RSEWYREGRVPLHTLRANVDYATSEAHTTYGVLGLKVVWYTEGNISKSPEHESKQRKAG  
 19pae RTEWYREGRVPLHTLRADIDYATYEAHTTYGIGVKVWIFKGEVIGGR--Q-EELKPVAP  
 21rp RTEWYIEGRMPLHTLRADIDYSTAEAITTYGIGVKVWYIKGEYKENKRYN-----  
 22syn RTEWYREGRVPLHTLRADIDYARTALTTYGILGKVVWIFKGEVIGPQEAIAVAPPSPQR  
 23tm RREWYLRGLPLQKIKAIIDYGTATATWTKYGTIGIKVWYIKGDADI-----  
 24tp RTEEIKERTPLHTLRADIDYGFAEATTYGIGVKVWYISGMMYNECKRDVGSLLR--  
 25vib RSEWYREGRVPLHTLRADIDYATSSAHTTYGIGVKVWIFKGEILGGMAAATAEAPKAD  
 26xyf RSEWYREGRVPLHTLRADIDYGFSEAHTTYGIGVKVWYIKGDIQVGGKRELAAPAGAD

```

01aae -----
04bb PFR----ERGDAVKNFDKTLNNREKANEKQTRLDDKKDGLSKDEVDLLNKKKFSASFSKE
05bs -----
06cj KKP----RRT---RRGK-----
07cp -----
08ct A-----
09der DER----RPEGGDRPARRRPTARRRTGGE-----
10ec PKK----QQ----R-KGRK-----
11hi KPK----KAP----RGKGRK-----
12hp REP----RRS---RRGRQ-----
13mg KKQ----LKTPTIKKTNSVIAKQKLTGSDIETASLKALTD-NNQNHE-----
15mp KKQPNRQNNQPRHFNQGGVLSANKLTGSDVETSSIQALTKPNKEDKQ-----
17mtu RPR----RERPSGTRPRRSASGTTATGTDAGRAAGGEEAAPDAAPVEAQSTES-----
18nm RRN----AAA---N-----
19pae APR----KK----AAR-----
21rp -----
22syn -----RKSRRQQFDDRSQDG-----
23tm -----
24tp --R----SRRESGQKSDELVRDERTHAERG-----
25vib KPK----KQ----R-KGRK-----
26xyf GDR----NAD---RSSRRSREVR-----

```

```

01aae -----MGGKDIDRLIEE
04bb RDDSNEQDIGG-----
05bs -----
06cj -----
07cp -----M
08ct -----M
09der -----MALTF
10ec -----AH--
11hi -----MSN--
12hp -----
13mg -----MNDQKTTNTGLLTSTLKT--KPKHNLKPSSEAIKKAWSKKEGHYKN
15mp -----MTDQNKANQNGGLQTTNLQAHAQRKHNLRPSSEGIKKAWSKKEGGGHN
17mtu -----MAEQPAGQAGTTDNRDARGDREGRRR
18nm -----MAKH-
19pae -----MANN-
21rp -----MSK--
22syn -----MAKRR
23tm -----
24tp -----MDRHR
25vib -----MAKE-
26xyf -----MAEER

```

```

01aae RRKQQNIEEILP-ELQLEERLIYANRTARVTKGRRFSFSTLVVVGDRKGHVGFHGKAK
04bb ---MVDVHAQR---KQIEKLISLNRVTKVVGGRRFSAAFMVVGDGEGHVWGFGKAN
05bs -----MRRIDPSKLEELERLVTVNRVAKVVGGRRFRAALVVVGDKNGHVGFGTGKAQ
06cj -----MEKYNREEFEVIVDIGRVTKVVGGRRFRTALVIVGNRKLGVGVGYKAK
07cp S-LSKNSHKED---QLEEKVLVNRCKSVKVGGRKFSFSAALILVGDGKGLGVGFYKAK
08ct T-LSRNSHKED---QLEEKVLVNRCKSVKVGGRKFSFSAALILVGDGKGLGVGFYKAK
09der N-----RRNDRTESSEFEKMLFVNRTSKTYQGGRRFRFAALVILGDRNGRVGMGIGKAK
10ec -----IEKQAGELQEKLIAVNRVSKTVKGRIFSFALTALTVVGDGNGRVGFYKAK
11hi -----IEKQVQELQEKLIAVNRVSKTVKGRIMSFTALTALTVVGDGNGRVGFYKAK
12hp ---MTERKGMEEINREEFQEVVVNIGRVTKVVGGRRFRAALVVVGDKNGLVGFGLGKAK
13mg KRQKHNFN---KSEFEERIVKLKRIKTTKGRNMRFSVLVVGNKKGKGVGYIAKAL
15mp RNNQNRFRQKPAFKSEFEERIVKLKRIKTTKGRNMRFSVLVVGNKKGKGVGYIAKAL
17mtu D-SGRGSRERDGEKSNYLERVVAINRVSKVVGGRRFSAALVIVGDGNGMVGVGYKAK
18nm -----EIEERGDGLIEKMVAVNRVTKVVGGRIMAFSALTALTVVGDGNGRIGMGKGSK
19pae -----EQKRDGEYIEKLVQVNRVAKTVKGRIFAFTALTALTVVGDGKGRVGFGRGKAK
21rp -----VKKNEEALSEVLVDVNRVTKVVGGRRFSAAYVVVGDKAGRVGAGHGKAK
22syn ---KTSREKK-EDTNWQERVQIRRVSKVVGKGLSFRAIVVVGNETGQVGVGVGKAG
23tm --METQGVMEIQYEEFEKIIEIRRTSKVTGKGNLSFRVVAIVGNKKGKVLGIGKAK
24tp D-FGKDRLRDK---EFTEKLIKLNRTAKVVGGRRFSAALTVVGDGKGRVGFYKAG
25vib -----QQVQANDLQEKLIAVNRVSKTVKGRIMSFTALTALTVVGDGNGRVGFYKAK
26xyf SQRNRDRSREEKIDDGMIKLVAVNRVSKTVKGRQFTFTALTALTVGNEGSVGFYKAK

```

```

01aae EVTLAIAKGIEKAKKNVIRVIVDG-TVPHDVI GTFGATKIIIVLPARRGTGVVAGGAAP
04bb DASDAIKKSLTSARKNLRFPV-IRKGTLPHEVIGCFKKAKVLKPAHTGTGVIAGGPVRA
05bs EVPEAIRKAVEDAKKNLIEVPMVGT-TIPHEIIGRFAGNILLKPASEGTVIAGGPVRA
06cj EVPDAIRKAVDDAFKNIVEVKTGS-TIAHDVEKYNASRILLKPASEGTVIAGGSTRP
07cp ELTDAIRKGGEAAKKNLMKIEALEDGSIPEHVLVNHDAQLLLKPAKPGTGIVAGSRIRL
08ct ELTDAIRKGGDAARKNLVINSLEGGSIPEHVLVNHDAQLLLKPAKPGTGIVAGSRIRL
09der EVPVAIEKAKAVARKNMTVPVENG-TIPHEIVGENSTRVLLKPAKPGTGIVAGTVPRS
10ec EVPAAIQKAMEKARRNMINVALNNG-TLQHPVKGVHTGSRVFMQPASEGTVIAGGAMRA
11hi EVPAAIQKAMEKARRNMINVALNEG-TLQHPVKGVHTGSRVFMQPASEGTVIAGGAMRA
12hp EVPDAIKKAVDDAFKNLHVITIKGT-TIAHDIEHKYNASRILLKPASEGTVIAGGSTRP
13mg EVPLAIIKKAIIKKAHNSIHTVEIHKG-SIYHEVIGRKGASKVLLKPAPLGTGIIAGGAIRA
15mp EVPNAIKKAIIKKAHNSLHTIEIHKG-SIYHEVIGRSGASRVLLKPAKPGTGIIAGGAIRA
17mtu EVPAAIKAGVEEARKSFVRPLIGG-TITHPVQGEAAGVLLRPAKPGTGIIAGGAARA
18nm EVPVAVQKAMDQARRSMIKVPLKNG-TIHHEVIGRHGATKVFVMPQAKESGVKAGGPMRL
19pae EVPAAIQKAMEAARRNMIQVDLNGT-TLQYPTKSAHGASKVVMQPASEGTVIAGGAMRA
21rp EVNEARGKAKQAARKRMKVPLYQNRTHHDVVGKSGAAKVILLRAKAGTGVIAAGSMRA
22syn DVIGARVKGVADGKKLIEVPLTKSNSITHITNGVSGGAKVVVRPAAPGTGVIAGGAVRT
23tm EVPEAIRKAIKKAHNSIHTVEIHKG-TIPHEVIGRQDASKVLLKPAKPGTGIIAGGTVRA
24tp DVSEAIRKSVARAKSMVLF-LKDGITIPHEVQAKFKGSLVLLRPAKPGTGIIAGGTVRA
25vib EVPAAIQKAMEKARRSMVTIALNEG-TLHHPVKGRHSGSKVVMQPAEAGTGVIAAGAMRA
26xyf EVPVAIQSMEYARKTMANVSLNNG-TLWHPVKANHGAACVFMKPASEGTVIAGGAMRA

```

01aae VFELAGYTDVLTIIIGSTNPNDNVRAVFDALLQLKTPAQIAEERGLPVEEILRRYRYAL  
04bb VMEALGVHDILSKSLGSNNMNVKATFKAFDVLNAEKVAEMRGKTLKTLWG-----  
05bs VLELAGVADILSKSLGSNTPIINMIRATLQGLSELKRAEDVAKLRGKSVVEELG-----  
06cj IVELAGIKDILTKSLGSNNANVVRATIKALTMLKG-----  
07cp ILEMAGIKDIAKSLGSNNPMNQVKAFAKALTGLSPKDLLR-RGAAIND-----  
08ct ILEMAGVKDIAKSLGSNNPMNQVKAFAKALTGLSKDDIMK-RRAVIND-----  
09der IELAGITNMLSKELGSNNKVNAYAVFDGFKNLRTAKQVRNLRGTEVPRSLSSDSPAGR  
10ec VLEVAGVHNVLAKEYGSTNPINNVVRATIDGLENMNSPEMVAARKGKSVVEILGK-----  
11hi VLEVAGVHNVLAKEYGSTNPINNVVRATIDALANMKSPMVAARKGKTVDEILG-----  
12hp IVELAGIKDILTKSLGSNNPNVVRATFDALAKIKA-----  
13mg IVELAGFSDIYTKNLGRNTPINMIHATMDGILKQLSPKKVALLRNKPISDL-----  
15mp IIELAGYSDIYTKNLGRNTPINMIHATMDGILKQLSPRRVAILRNKNLNEL-----  
17mtu VLECAGVHDILAKSLGSNAINVVHATVAALKLLQRPVEEVAARRGLTIEDVAPAGMLKAR  
18nm VFDAMGIHNSIAKVHGSTNPINVRATLDGLSKLHTPADIAAKRGLTIEDVILGVNHG---  
19pae VLEVAGVQNVLAKEYGSTNPVNVYATFKGLKNMQAPEAAAKRGKSVVEIL-----  
21rp IFDSLGVHDIAKSLGSNTVYAMISATFDALNKLASPKSIAIRRDKKVHEISIKSYIQVN  
22syn VLELAGVKNILAKQLGSNNPLNNARAAINALETLRTFSEVAEERGVSEHLYT-----  
23tm VVELAGIQNITLTKSLGSTNPLNLAATMNGLKNLLDPKRAKLRDISVEEVFKGVRRENN  
24tp IMEVAGATDVLKSLGSNSAINVVVRATFGAVALMDARKLARERKGLVDMWG-----  
25vib VLEVAGVHNVLAKEYGSTNPINIVRATIDALVDVKSPMVAARKGLTVEAISE-----  
26xyf VLEAVGVKDVLAKEISGRNPINLVRATLKGLEDMSPTIHALKRGKNVGRGFSHGS-----

01aae PKMNKHYYPWRGIYEQTMPKGPVPPREIPDPKYGDVLVQKLINKVMKDGKKSVAEWI  
04bb -----MSRKNKKIKKKVFDTRYNSRIVAKFANRMMYDGKKSISESI  
05bs -----MPKGPVAKRDVLPDIYNSKLVSRLINKMMIDGKKGKPTI  
06cj -----MRRRKAPVREVLDPDIYGNKVITKFINSLMYDGKKSATTI  
07cp -----MSRRHSAEKRDIPGDIYGSVILEKFINKVMHGGKSVARKI  
08ct -----MSRRHAAEKVPIGDPVYGSVVLERFINVKMLHGKKSIAKI  
09der SATTEAGEGVADTGGMQ-MARRRRAEVRPVQDVLVYQDVLVSAMINIRMDGKKNLASRI  
10ec -----MPRRRVIGQRKILPDPKFGSELAKFVNILMVDGKKSIAESI  
11hi -----MPRRRSVEARKILPDPKFGSELAKFINVIMVDGKKSIAESI  
12hp -----MRRRKAPVREVLDPDIYGNKVITKFINKMMFDGKKSVAEKI  
13mg -----M-RKNRALKRTVLDPDPVFNNTLVTRIINVIMKDGKGLAQRI  
15mp -----M-RKNRAPKRTVLDPDPVFNNTLVTRIINVIMEDGKGLAQRI  
17mtu RKSEALAAVLPDRTI--MPKGPAPKRPVNDPVYGSQVLTVLQNVLLKGKKSIAERI  
18nm -----MPRRREVPKRDVLPDPKFGSELAKFVNILMVDGKKSIAERI  
19pae -----MPRRRVAAREVLADPKYGSQILAKFMHVMSEGGKKAERI  
21rp E-----MSRRHVAEKRIILPDMKYNISILSRFINIMKAGKKAERI  
22syn -----MSRGNVKKRPVPPDPVYNSTLSMTIRVMRSKKSGLASSI  
23tm A-----MRRRRAEKQIPDPVFGDVLVAKLINRMVMDGKKTIAQKI  
24tp -----MGRKRVRSRVPPDPDARYNSVVLAKFICRMMLAGKKAATAVGI  
25vib -----MPRRRVIGQRKILPDPKFGSELAKFVNILMVDGKKSIAESI  
26xyf -----MSRKGSTPQRNVLPDPKYGSETIARFINMVMKSGKKSVAEKI

01aae VYTALAEAAKEVNMH-PVELLHKVIEKLKPEWEVRPRRVGGATYQVPIEVPERRQISLAI  
04bb LYSSIDLLADKLEESDKMAVFYKALDNKPLVEVRSRRVGGATYQVPIEVEERREALAM  
05bs LYKSFDIKERTGND-AMEVFEQALKNIMPVLEVKARRVGGANYQVPIEVPERRTTGL  
06cj MYGALEAIDKGGEGKIDIFNDATENIKPLLEVKSRRVGGATYQVPIEVPARQQALAI  
07cp VYSALERFQKLLNLENVLEGFGEALENAKPILEVRSRRVGGATYQVPIEVASERRNCLAM  
08ct VYGALERFAKRLGLENPLEGFGEALENAKPILEVRSRRVGGATYQVPIEVAPDRRSCLAM  
09der FYGACRLVQERTGQE-PLKVFKQAYDNVVKPRVEVRSRRVGGATYQVPIEVGPRRQQLSTL  
10ec VYSALETLAQRSGKS-ELEAFEEALENVNRPTEVKSRRVGGATYQVPIEVRPVRNNALAM  
11hi VYGALETLAQRTGKE-PLEAFEEALENVNRPTEVKSRRVGGATYQVPIEVRPVRNNALAM  
12hp IYKAFNKKIEKSGEK-GIEVFEKALERVRPLVEVRSRRVGGATYQVPIEVASRRQQLSLI  
13mg LYGALEIIKERTNQ-PLTVFEKAVDNVMPRLKVRRIAGSNVQVPTVPPDRRIALAL  
15mp LYGAFDLEIQRTEKE-PLTVFERAVGNVMPRLKVRRIAGSNVQVPTVPPDRRIALAL  
17mtu VYGALEQARDKTGTD-PVITLKRALDNVKALEVRRSRRVGGATYQVPIEVRPDRSTTLAL  
18nm VYGALEQIEKTKGV-AIEVFNEAIAKAPIVEVKSRRVGGANYQVPIEVRPVRNNALAM  
19pae VYGALDKVKEGKAD-PLETFEKALDAIAPLEVKSRRVGGATYQVPIEVRPVRNNALAM  
21rp VYSALNKIEKHSVD-PYQTFNNAMHNKPHLEVTSVRVGGANYQVPTVHDERRGYALAS  
22syn VYNALASVGEKTGED-PLEVFEKATKNLTPLVEVKARRVGGATYQVPIEVRPARGTALAL  
23tm VYGAFDIIREKTKD-PLEVFRQAVENVKPILEVRSRRVGGATYQVPIEVEPRRTSLAL  
24tp MYDCLERIQQRTGE-EPLPVFTKALENVKPAVEVKSRRVGGATYQVPIEIRETRREALGM  
25vib VYTALDTMAEKSGKD-HLAVFEEALENVRAVEVKSRRVGGATYQVPIEVRPVRNNALAM  
26xyf VYGAMNVIKKN-SN-AIELLQKALDNVSPAVEVKSRRVGGATYQVPIEVRASRRMALAM

01aae KWLVAARERPRGRGQYTMIERLKAELLDALNERGGAYKKKEETHRMAHANMVFSHFRW-  
04bb KWIIFAARKS----SGRSMKEKLSNELLNAYNSTGAFFKKKEDTHRMAEANKAFHYRW-  
05bs RWLVNYARLR----GEKTMEERLANEILDAANNTGAAVKKKREDTHKMAEANKAFHYRW-  
06cj RWIISFARKR----SERTMIDKLAAELDDAANSKGASFKKEDTYKMAEANKAFHYRW-  
07cp QWIIKHARKS----PGKSMEVGLATELIDCFNKQGATIKKREDTHRMAEANKAFHYRW-  
08ct QWIIKHARKS----PGKSMEVGLATELIDCFNKQGATIKKREDTHRMAEANKAFHYRW-  
09der RWMISAVDGR----PERTAIERLAGEIMDAAGRGGAIAKKKEDVHRMAEANKAFHYRW-  
10ec RWIVEAARKR----GDKSMALRLANELSDAENKGTAVKKREDVHRMAEANKAFHYRW-  
11hi RWIVEAARKR----GDKSMALRLANELSDAENKGAIVKKREDVHRMAEANKAFHYRW-  
12hp RWILEATRRK----NERMMVDRLANELMDASDKGAFFKKKEDVHKMAEANKAFHYRW-  
13mg RWIVIFANKR----NEKTMLEVRANEIIDAFFNTGASVKKKDDTHKMAEANKAFHYRW-  
15mp RWIAMFARKR----HEKTMLEKIANEIDASNNTGAAIKKEDTHKMAEANKAFHYRW-  
17mtu RWLVGYSRQR----REKTMIERLANEILDAENKGTAVKKREDTHKMAEANKAFHYRW-  
18nm RWVRDAARKR----CEKSMALRLAGELIDASEGRGGALKKREEVHRMAEANKAFSHFRF-  
19pae RWLVDFARKR----CEKSMALRLAGELIDASEGKGAIVKKREDVHRMAEANKAFSHFRF-  
21rp RWIINAASKR----SEKMMIDKLAELFEASNNRGVAIKKEDTHKMAEANKAFSHFSPK  
22syn RWLVHFSRAR----GGRTMESKLANEIMDAANETGAAIKKREDTHRMAEANKAFHYRW-  
23tm RWIVEAARAK----KGRPMKEKLAEEIIAAYNTGTAIKKEDTHRMAEANKAFHYRW-  
24tp RWIIGAARRR----SGRGMERLAAEILDAYHSTGAFKKREDTHRMAEANKAFHYRW-  
25vib RWIVEAARKR----CEKSMALRLAAEMLDAENKGAIVKKREDVHRMAEANKAFHYRW-  
26xyf RWLIDSSRRK----GENSMPHKLAELLDASESRGGAIAKKREDTHRMAEANKAFHYRW-

01aae -----MSAVDPIADMFSAIKNAIMRRDDFLYPSS  
04bb -----MAITYSIGDMLTKLRNASRVGHGSVDLKMS  
05bs -----MVMTDPIADMLTRIRNANMVRHEKLEIPAS  
06cj -----MINDIISDSLTRIRNAGMRKLETTKLLHS  
07cp -----MGMTSDSIADLLTRIRNALMAEHLVDVEHS  
08ct -----MGMTSDSIANLLTRIRNALMAEHLVDIEHS  
09der -----MLSDPIADMLTRIRNATRTHKETVDIPAS  
10ec SLRSFSHQAGASSKQPALGYLN-----MSMQDPIADMLTRIRNGQAANKAAVTMPSS  
11hi -----MSMQDPIADMLTRIRNGQAANKVAINMPSS  
12hp -----MVNDIADSLTRLRNASMRRLFTQLYYA  
13mg -----MIINKVP-KAHFDPVSDLFTKINNARKAKLLTVTITIAS  
15mp -----MITTTPIKAHFDPPVADLLTKINNARKAKLMTVTITIAS  
17mtu -----MTMTDPIADFLTRLRNANSAYHDEVSLPHS  
18nm -----MSMHPDISDMLTRIRNAQRANKAAVAMPSS  
19pae -----MSMQDPLADMLTRIRNAQMAEKTVVSMPSS  
21rp KMK-----MSMTDNVADMLTRIRNAYSKLINVSFPSS  
22syn -----MASTDTISDMLTRIRNACAVRHSTTQVPTT  
23tm -----MLTRIRNANMVFEKTDIPAS  
24tp -----MGVSDPVADMLTKIRNARAGHEKVDVPSS  
25vib -----MSMQDPIADMLTRIRNGQAANKVAVKMPSS  
26xyf -----MSMTDPIADMLVRINNAASVGKPNVRFPFS

01aae KLKERILDVLKKEGFIQDWEALKGEYEEYKKMKELAEKSPNPKMKRYLKQLEEYKNGT  
04bb MNKSIILNLKEEGYIKDFNFLE-----KKG  
05bs KLKREIPEILKREGFIRDVEFVE-----DSK  
06cj KVVVALVGIFQAKGYIESFNVIE-----EDK  
07cp KMREAIKILKHKGFVAHYLVKE-----ENR  
08ct KMLEAIVRILKQHGFAHFLVKE-----ENR  
09der KFKEELAKLLVREGYVQGVERTPE-----GQK  
10ec KLKVAIANVLKEEGFIEDFKVEG-----DT  
11hi KLKVAIANVLAEGYIESVKVLE-----GA  
12hp KIVVSILEIFKEKGFIDFNVKD-----KDK  
13mg KLKIAILEILKEGYLANQVLENK-----TKT  
15mp KLKIAILEILKVEGYLANFQVLENK-----SKT  
17mtu KLKANIAQILKNEGYISDFRTED-----ARV  
18nm KLKCAIAKVLKEEGYIEDFAVSS-----DV  
19pae KLKAAVAKVLDEGYIADFQISS-----EV  
21rp KIKTSILDVLKKEGYIKDYVTTQ-----KNN  
22syn KMTLSIAKVLKSEGFIEDYSETG-----EGI  
23tm NLKKKI CEILKREGFIADYKYIE-----DGK  
24tp KLKVEVVVKLKTGYIRNFRKVE-----EDG  
25vib KLKVAIAALKAEGYIADFAVEG-----EV  
26xyf RVKLAIALVLKHEGYIFDAKVIQ-----DNS

01aae QYPIKIYLYLDPKKR-KSAITNIVKVSKGRRVYAGVRTMPYVKRGLGIAIVSTDAGVM  
04bb IAFIRVLLKYD---NKRNPVINKIDAISTGRKIYSSYRNMPRIKNGYILIISSSQGVI  
05bs QGIIRVSLKY---GQNN-ERVITGLKRISKPLRVYAKSNEVPRVLNGLGIAIISTSQGVL  
06cj KKFINVVLKYD---EGKSVINELKRISKPGRRVYKGDDEIKRFPKNGYGTIVVSTSHGVL  
07cp KRAMRVFLQYS---DDRKPVIHQLRVSKPSRRVYVSAKIPYVFGNMGISVLSTSQGVM  
08ct KRLMRVFLRYG---EDRRPVIHALKRVSKPSRRVYVSAKIPYVFGNMGIAVLTSTQGV  
09der FDLVRVTLKYG---EKREQVIKHIERVSRPGRRAYVSAENLPRVQKGLGVAVVSTSRGLL  
10ec KPELEITLKY---FQKAVVESIQRVSRPGLRIYKRKDELPKVMAGLGVAVVSTSKGVM  
11hi KPELEITLKY---FQKPVVESIQRVSRPGLRIYKRKDELPKVMAGLGVAVVSTSKGVM  
12hp KQSVYVQLAYD---EKGHSKISEVKRLSKPGRRVYKQKDELKRFKNGYGVIVVSTSKGVI  
13mg KKLVSFTLKYT---QRRICISINGVKIQSKPGLRIYRSFEKPLVLNGLGIAIISTSDGVM  
15mp KRIVTFNLKYT---QRRIPSINGVKIQSKPGLRIYRPFELPLVLNGLGIAIISTSDGVM  
17mtu GKSLVILQKY---GPSRERSIAGLRVSKPGLRVYAKSTNLPVGLGGLGVAIISTSSGLL  
18nm KSILEIQLKY---YAGRPVIEQIKRVSRPGLRIYKASSEIPSVMMGLGIAIVSTSKGVM  
19pae KPQLSIELKY---FEGKPVIEEVKRISRPGLRQYKVEQLPKVKGGLGVSVSTNKGVM  
21rp ISYTEVALKYS---VNGEASICEIHRVSKPGKRVYSAIKDLKGYNNMGYIILSTPYGVM  
22syn NKMLVTLKYK---GKTRQPLINTLQRVSKPGLRVYSPSKKIPRVLGGIGIAIVSTSHGIM  
23tm QGILRVLYKYKGRKNRERVIGHIVRVSHAGRRIVYVDKDHIPKVMNGLGIAILTTSKGV  
24tp SGCIRVFLKYD---DNETSIVHIERISTPGRRVYSGYKTLRAVYNGYGTILVSTSLGVT  
25vib KAELEITLKY---FQAKPVIEQIKRVSRPGLRVYKDELPSVMMGLGVAVVSTSKGLM  
26xyf KSDIEIQLKY---FEGRPVIRILKRVSRGLRKYCGKAELPKVLGGLGISIISTSKGIM

01aae TDHEARRMRKGGEVIAFW-----MIQKLKDFKLTPENTYYATG  
04bb TGKEAKDKKIGGELICSVW-----MKKSNSFNVLNLSMTGT  
05bs TDKEARAKQAGGEVLAYV-----LAQVQYYGTG  
06cj ANDEAYKAGVGGEILCTIW-----MATTYATG  
07cp EGSLSAKSNIGGELLCLVW-----MAKSTIQES--VATG  
08ct EGSVARAKNVGGELLCLVW-----MIQES--VATG  
09der PDREARQQVGGEVVCVLW-----MAIQQPEQYYGTG  
10ec TDRAARQAGLGGEIICYVA-----MAENQYYGTG  
11hi TDRAARQAGLGGEIICYVA-----MAENQNYGTG  
12hp TNEEAYRQNVGGEVLCSIW-----MRKIYATG  
13mg TDKVARLKKIGGEILAYV-----MDKKSFYGLG  
15mp TDKVARLKKIGGEILAYV-----MEKQSYYYGLG  
17mtu TDRQAARQVGGEVLAYV-----MTETTPAPQTPAAPAGPAQSFVLERPIQTVG  
18nm TDRKARSQVGGEILLCIVA-----MNGKYYYGTG  
19pae TDRAARAAGVGGEVLCTVF-----MSATQNYGTG  
21rp SDREAHIKNVGGEVICKVMTTELKIKTEKVVKQLTKESLSVLKIPKEKIDSVSIFYATG  
22syn TDREARQQGIGGEILCIW-----MQANDSSNKVYVYGTG  
23tm TDKEARPVGVGGEVIAFW-----MAEVVGYGTG  
24tp TGRHAREQRVGGEILCKVW-----VKNLGIGGTG  
25vib SDRAARKAGLGGEIICYVA-----MAENQYYGTG  
26xyf IDSKARES VGGEVLCFVA-----MPITQNYGTG

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01aae | RRKEAVARVLLRDRPNTFIVGSDKTNKEYDLREYVQRETLYNKIMYPFKVTG--HEGKF     |
| 04bb  | RRKSSVARVYIREGS--GNIKVNNRDPDSY----IQ--LENLRMTALSPVLVTN--TLGKY   |
| 05bs  | RRKSSVARVRLVPG--GRIVVNNREISEH----IP--SAALIEDIKPLTLTE--TAGTY     |
| 06cj  | RRKTAIAKVVWPKGS--GKISVNGVDLNTW----LGGHEAIKLVVQPLLVTK--QETSM     |
| 07cp  | RRKQAVSSVRLRPGS--GKIDVNGKSFEDY----FP--LEIQRTTILSPKKITE--DQSQY   |
| 08ct  | RRKQAVSSVRLRSGN--GKIDVNGKTELEY----FP--LEVQRATILAPLRMLG--DVNSF   |
| 09der | RRKAAVARVFLRPG--GKIVVNGKEFTQ-----YFRGLLRVNALQGFRETGTACRF        |
| 10ec  | RRKSSAARVFIKPGN--GKIVINQRSLEQY----FG--RETARMVVRQPLELVD--MVEKL   |
| 11hi  | RRKSSSARVFIKPGS--GKITINQRELDVY----FG--RETARMVVRQPLELVX--LTDKL   |
| 12hp  | RRKTAIAKVVWLPKG--GELSINEQSLNQW----LGGHEAIKMKVMQPLLT--QEQSV      |
| 13mg  | RRKSSSTAKVYLYQSKDKGKITINHRNPSDY----FP--NKLVIQDMEQPLELTK--LKDNF  |
| 15mp  | RRKSSSAKVYLTPTQDKGKITVNRDPSEY----FP--NKLVIQDMEQPLELTD--LKKNF    |
| 17mtu | RRKEAVVRVRLVPGT--GKFDLNGRSLEDY----FP--NKVHQQLIKAPLVTVD--RVESF   |
| 18nm  | RRKSSVARVFLIKGT--GQIIVNGRPVDEF----FA--RETSRMVVRQPLVLT--NAESF    |
| 19pae | RRKTATARVFLRPGT--GKISINNRGLDQF----FG--RETARMVVRQPLELTE--TVEKF   |
| 21rp  | RRKNAIARVWLKVGK--GKIVVNNKILNQY----FP--SETYVKTILQPFILTK--TIDQY   |
| 22syn | RRKAAIARVRLVPGQ--GEVIVNGKPGEY----FNRIANYIQLSKAPLETG--LEGEY      |
| 23tm  | RRKTAVARVYLRPG--GKVIVNGKEYESL----NDYFKNPAWTKHAIPELVNTLTKGF      |
| 24tp  | RRKTAVARVCIRMGN--GNVTVNRDVGAY----FP--TAEQLRRVREPLFATA--NERRY    |
| 25vib | RRKSSAARVFIKPGS--GNIVINKRSLEEY----FG--RPTSCMVVKQPLELVD--MVEKL   |
| 26xyf | RRKSSKARIFLRKGS--GNITVNGRPLDEF----FG--RETARMIVRQPLELTK--NVANF   |
|       |                                                                 |
| 01aae | GIYATVRGGGISAQAEAIMYGVAKALLQHNPDLRPTLKKAGLLTRDAREKERRYQGMGA     |
| 04bb  | DLVINVYGGGISGGQGAIRHGISRALFKLDESNMILRSNGFLTDRSRKVERKKFGQKKA     |
| 05bs  | DVLNVVHGGGLSGQGAIRHGIRARALLADPEYRTTLKRAGLLTRDARMKERRYGLKGA      |
| 06cj  | DIKATTLGGGYSQAQAEALRHGISRALAAMDADFRALLKPKGLLTRDSRTVERKKYGRKA    |
| 07cp  | DLIIRVSGGGIQQGVIAITRLGLARALLKENEENRQDLKSCGFLTDRPKKERRYGHKKA     |
| 08ct  | DLIIRVSGGGVQQGVIAITRLGLARAVLQEKEDMKQELKAQGFLTRDPKKERRYGRKKA     |
| 09der | DAVITVTGGGSGQADAIAKLGARALLKVNPDFRAQMKPKGLLTRDPREVERKKYGLKKA     |
| 10ec  | DLYITVKGGGISGGQGAIRHGITRALMEYDESLSRELKAGFVTRDARQVERKKVGLRKA     |
| 11hi  | DLYITVKGGGISGGQGAIRHGITRALMEYDETLRALRAAGFVTRDARRVERKKVGLHKA     |
| 12hp  | DIKAVVFGGGYSQAQAEALRHGISKALNAYDIAFRAILKPKGLLTRDSRVVERKKYGRKA    |
| 13mg  | DINVVVGGGGFTGGAGAIRLGIVRALIKFNPDLKLLTKKLLTRDKRAKERRYGLGYGA      |
| 15mp  | DINVVVGGGGFTGGAGAIRLGIVRALLQFNPDLKLLTKKLLTRDKRVKERRYGLGYGA      |
| 17mtu | DIFAHLLGGGGSGGAGALRGARALLVSPEDRPALKKAGFLTDRPRATERKKYGLKKA       |
| 18nm  | DIKVVVGGGGTGGQGAIRHGITRALIDFDAALKPALSQAGFVTRDAREVERKKVGLRKA     |
| 19pae | DIFVTVGGGGSGGQGAIRHGITRALIEYDETLRSSLRKAGYVTRDAREVERKKVGLRKA     |
| 21rp  | DVICTVKGGGISGGKAILHGISKALDSPAFCFHAILRKGLLTRDSRVVERKKYGRKKA      |
| 22syn | NILVNAHGGGLTGQADAVKGVARALCQLSPENRQPLKAEGYLTDRPRAKERKKYGLHKA     |
| 23tm  | DLVIRVNGGGLSGGQGAIRLGARALLQYDQNLRLPVLLKYKMLTRDPREVERKKYGLKKA    |
| 24tp  | DVIVNVYGGGLDGGAGACAHGIARALVRADASNQASLRAGGLLTRDSRMVERKKYQGRGA    |
| 25vib | DLYITVKGGGISGGQGAIRHGITRALMEYDESRLPVLRAAGYVTRDARRVERKKVGLRKA    |
| 26xyf | DILITATGGGTGGQGAIRLGARALVEYDASLPELRKAGFMTDRDAREVERKKVGLHKA      |
|       |                                                                 |
| 01aae | RAKYRWSKR--MSEYLKALFELK-----ELRQK-----FIFTLFMFVIYRLGSHIP        |
| 04bb  | RKSFQFSKR--MKE--LFLSLFTVK-----DLRKN-----FLFTLVFLFRVGSYLP        |
| 05bs  | RRAPQFSKR--LFKTI SNFMRVS-----DIRNK-----IIFTLMLVFRIGAFIPV        |
| 06cj  | RRSPQFSKR-----MN-----RALTNK-----ILITLAFLFAYRVLAYVPV             |
| 07cp  | RKSFQFSKR--MFRPYMTTLRQFFL-----ITELRQK-----LFYTFALLTACRVGVFIPV   |
| 08ct  | RKSFQFSKR-----MATLRQVFS-----ISELRQK-----IFFTSLALLCRIGVFI        |
| 09der | RRAPQFSKR--MLRAFRDAFRIP-----DLRQK-----IVFTLLLLAVYRLGGAIPT       |
| 10ec  | RRRPQFSKR--MAKQGLDFQSAKGG-----LGELKRR-----LLFVIGALIVFRIGSFIP    |
| 11hi  | RRRPQFSKR--MAKQPGYQARSTNSG-----KGELKSR-----LLFVLGALIVYRIGSFIPV  |
| 12hp  | RRSPQFSKR-----MN-----KAIASK-----ILITLGLFLYRVLAYIP               |
| 13mg  | RRAPQFTKR--MQT--VSSPKQKLNFGQRLTLQNRDMVSLVTLVLLILFRVLAIPL        |
| 15mp  | RRAPQFTKR--MQAKPTTAVKQKQNFQGRFLTLLRNDRFMISFLITVLLVFRVLAIPL      |
| 17mtu | RRAPQFSKR--VLSAFISSLRTV-----DLRRK-----ILFTLGIVILYRVGAALPS       |
| 18nm  | RRAKQFSKR--MANQQTSSGSS--K--FGDLKRR-----LLFLFGALIVFRIGAHIPV      |
| 19pae | RRRPQFSKR--MAKQGALSALS--NGG-----LSELWAR-----LRFLFLAIIVYRIGAHIPV |
| 21rp  | RKKTQFSKR--MGQNFSSKSSN-----DLVNR-----IIFTLFMLICRFGSFIP          |
| 22syn | RKAPQFSKR--MVSRDKAPSQETFLQMAQAAGLRGR-----LLITIGLLILVRVGFIPV     |
| 23tm  | RRAPQFSKR--MWQAFKNAFKIP-----ELRDR-----IIFTLALIVFRMGIYIPV        |
| 24tp  | RRRPQFSKR--MKQGVFAAVFRIR-----ELRAR-----IFFTSLVTVFRFGSVLTV       |
| 25vib | RRRPQFSKR--MARKPGQDFRSQSG-----LSELKSR-----LFFVIGALLVFRAGSFVPI   |
| 26xyf | RRATQFSKR--MSQPGGLGNLGRNFGK--SIELRHR-----LLFLFGALIVYRVGYVPV     |
|       |                                                                 |
| 01aae | PGINPEALRDFLK----AFEG--SVFALYDIFSGGNLGRLTVFALGVMPYISASIMMQL     |
| 04bb  | PGIDSVALKSIFYK----SQSD--FSIANFYDFFSGGAFSNFISFMLSIGPYISASIIVQL   |
| 05bs  | PYVNAEALQA-----SQSMGVFDLLN--TFGGGALYQSFIFANGITPYITASIIQL        |
| 06cj  | PGVNADVIAEFFN----NNQN--NALGLFNVFSGGAERFSIIISLGIMPYITASIIIMEL    |
| 07cp  | PGINGELAVAYFK-----QLLGGGQNLFLQDLIFSGGAFAMQTVIALGVVPYISASIIVQL   |
| 08ct  | PGINGDRAVAYFN-----QLLGGSSQNLFLQDLIFSGGAFAMQTVIALGVVPYISASIIVQL  |
| 09der | PGVNSAALSQ-----ANSGLFLGIS--MISGGNLQSFSIFALGVLPYITASIVIQL        |
| 10ec  | PGIDAVALAKLLE----QQRG--TIIEMFNMFGGALSASIFALGIMPYISASIIQL        |
| 11hi  | PGIDAVALAQLVE----QKQG--TIIEMFNMFGGALSASIFALGIMPYISASIVIQL       |
| 12hp  | PGVDLAALKAFFD-----SNSN--NALGLFNMFGGALSASIFALGIMPYITSSIIIMEL     |
| 13mg  | PGTRINESVLD-----RNSN--DFFSLFNLGGGGLNQLSLFAVGISPYISASIIIMQL      |
| 15mp  | PGIQVNQVTGLD-----QNSN--DFFSLFNLGGGGLNQLSLFAVGISPYISASIIIMQL     |
| 17mtu | PGVNFNQQCIKEASAGEAQIYSLIN--LFSGGALLKLTVFAGVMPYITASIIVQL         |
| 18nm  | PGVDAVALAKLYE----SAGN--GILGILNMFSGGSLERFSIFAIGIMPYISASIIVQL     |
| 19pae | PGINPDRLAALFR----QNEG--TILSLFNMFGGALSASIFALGIMPYISASIIIMQL      |
| 21rp  | PGIDSIALNSVAE----KNQF--GILGMFNMFGGSLGRMSIFALGIMPYITASIIIMQL     |
| 22syn | PDIDRQAFSQAIN--DN-----SVIGFLNIFTGGGLSTVGIFALGILPINASIIIMQL      |
| 23tm  | PGLNLEAWGEIFRRI--AETAGVAGILSFYDFTGGALSRSFVSFTMSVTPYITASIIQL     |
| 24tp  | PSVDPRALSAYFR--SQVRG--NAFADYMDFFVGGGAFSNFVFMGLVMPYISTQILMQL     |
| 25vib | PGIDAVALAELFE----QKQG--TIVEMFNMFGGALSASIFALGIMPYISASIVVQL       |
| 26xyf | PGVNPEAMLALMQ----KQGG--GIVDMFNMFGGALHRSFIFALNMPYISASIVVQL       |

01aae LTV-AIPSLQRLAKEEGDYGRYKINEYTKYTLTFVATVQSLGIAFWIRGQVS-PKGPV  
04bb LVY-SFPSSKKMQ--EGDGGGRQTKKYTKYTLTIVAAVQGYATSLYAKGIP-----GAVT  
05bs LQMDVVPKFTWSK-QGEVGRKLAQFTRYFTIVLGFIALGMSYGFNN-----LANGM  
06cj LAA-TFPNIGMKKK-ERD-SMQKYMQUIRYATVITLVQSIGVAIGLQSLHGRGGAGAIM  
07cp FLV-FMPALQREMRESSDQGRKRIGRLTRLFTVALAVIQSLLFAKFALRMN--LTIPGIV  
08ct LVV-FMPALQREMRESSDQGRKRIGRLTRLFTVALAVIQSLLFAKFALRMN--LVVPGIV  
09der LTT-TIPSLKLSK-EGEGRKKINQYTRYAIALGAVQALFFSLYIT-----SNPSF  
10ec LTV-VHPTLAEIKK-EGESGRKKISQYTRYGTLVLAIFQSIGIATGLPNMPG---MQGLV  
11hi LAT-VSPALAEKK-EGAAGQRKISKYTRYATVVFATIQAVAISTGLPNM-----LSGLV  
12hp LSA-TFPNLAEMKK-ERD-GMQKYMQUIRYATVITLVQSIGVAIGLQSLHGRGGAGAIM  
13mg LSTDLPPLSKLVN-SGEVGRKKIEMITRIITLPFALVQAFVIAQIATNAGT--GSSPIS  
15mp LSTDLPPLSKLVN-SGEVGRKKIEMITRIITLPFALVQAFVIAQIATNAGT--GSSPIS  
17mtu LTV-VIPRFEELRK-EGQAGQSKMTQYTRYALAILQATSVIVALAANGGL--LQCCS  
18nm ASE-ILPSLKALKK-EGEAGRKVITKYTRYGTLLAILQSLGVASFVFPQ-----GIV  
19pae MTA-ISPQLEQLKK-EGESGRKKISQYTRYGTLLAILQSLGVASFVFPQ-----GVA  
21rp MSV-AYKPLENKK-EGETGKRKINQLSRYLTVLLASFQAYGVALSLESMT-N-TGPVV  
22syn LTA-AIPALQREMRESSDQGRKKISQYTRYAFGWCIQGLGLTVGLLR-----PYA  
23tm LAS-VMPSLKEMLR-EGEGRKKIFAKYTRRLTLLIGGFQAFVFSFLAR-----SNPDM  
24tp SMI-VFPSLKVV--EDVGRRRVQFVTRVATVFCVLIQSSAVTVYANQIP-----CAIV  
25vib LTV-VHPALAEKK-EGEAGRKISQYTRYGTLLVATFQAIQIATGLPNM-----VNNLV  
26xyf ATH-IFPSLKAMQK-EGESGRKKITQYSRIGAVILAVIQGSIALLALQNTA-PGGAHV

01aae ENPGIS-----FILITVTLVAGTMFLVWIADRITEKGIGNGASLIIFAGIVANFP  
04bb IPFYR-----YIFVAILTVTTGTFTILLWFGEQINQRGVNGTSLIIFSGIVVRLQ  
05bs LIE-----KSGVSTYLIIALVLTGTAFLMWLGEQITRSHGVNGSIIIFAGIVSSIP  
06cj VEDLNM-----FIALCAISMLAGTMMLMWLGEQITRQGINGISLIIFAGIVSGIP  
07cp LPTLLSSKLFVGPWFIFYITTVVMTTGTLLLMWIGEIQSDKGINGISLIIALGILSSFP  
08ct LAPMLSKLFGVPVWFYLTTVVMTTGTLLLMWIGEIQSDKGINGISLIITLILGILSFP  
09der IAVGW----DPGLFTVLVMTLTVQAGIAFTMWIGERITEVGINGISLIITAGIIVAVP  
10ec INPGFA-----FYFTAVVSLVTGTMLMWLGEQITERGIGNGISIIIFAGIVAGLP  
11hi PNVGFS-----FYFTSVVSLVTGTMLMWLGEQITERGIGNGISILVFGGIVAGLP  
12hp ID-MQV-----FMIVSAFSMLTGTMLMWLGEQITRQGVNGISLIIFAGIVSGIP  
13mg LANSQS-----E----FIAFYIIMTAGTYMAVFLGDTISKKGVNGITLLILSGILSQPL  
15mp LKNGS-----D----FVAFYIIMTAGTYLTVFLGDTISKKGVNGITLLILSGILSQPL  
17mtu LDIIA-----DQSIFTLVVIVLMTGGAALVMMGELITERGIGNGMSLIFVGIARIP  
18nm VTSSFE-----FHVSTVSLVTGTMLMWLGEQITERGIGNGISLIITAGIASGIP  
19pae FSNDFG-----FYFVAVTTFVAGAMFMMWLGEQITERGIGNGMSLIFAGIVAGLP  
21rp ILAGFF-----FRVTVITLVVGTILLMWLGEQITRQGINGTSLIIFIGIISGVP  
22syn NNYGP-----LFIFQTVLITAGSMFVMWISLITERGIGNGASLIFVNIATPL  
23tm VAPGV-----NVLQFTVLS-TMSLAGTMFLWLGERITEGIGNGISILIFAGIVARYP  
24tp IQSYAV-----HLFVTMLTVTSGSMITLWLGEQITARGINGVSMIFSGIVARLP  
25vib VIDQTM-----FTLIATVSLVTGTMLMWLGEQITERGIGNGISILIFAGIVAGLP  
26xyf YAPGAA-----FVMTAIVALTAGTVFLMWVGEQVTERGIGNGMSLIFAGIVAGLP

01aae NAVIQFYE-KVKTG-----DIGPLTLLIILIAIIIVGIVVYQEAERRIPQYPR--  
04bb AALFNLFSQMQDPQ-----NVNPFVILIIISIFILVILIIIEYKQAMRIAIHYARA--  
05bs KTIQIYE-TQFVGSND-QLFIHIVKVALVIAILAVIVGVFIQVARKIAIQYAKG--  
06cj RAISGTVG-QINSQ-----EMNFLTAFALFALITIGVIIYIELGERRIPISYSRK--  
07cp SVLGSIVN-KLNLGSQD-SSDLGLISILALVVFVLTITLIIIEGVKIPVQYARR--  
08ct SVLGSIFN-KLNLGSQD-PSEFGIVSLLILCAVVFVLMATVLIIEGMRKIPVQHARR--  
09der REIAA----TAQLLRSE---QTTLLSILAFIAVILVTIAGIVVYIQAERRVPVYARARG  
10ec PAIAHTIE-QARQG-----DLHLVLLVAVLVAFTVVFVVERGQRRIVVNYAKR--  
11hi SAILQTIE-QARQG-----QMPLVLLIAAIVFAVTVFVVFVERGQRRIRVEYAKR--  
12hp SAISGTFN-LVNTG-----VINILMLIGIVLIVLATIFAIIVVELAERRIPISYARK--  
13mg QGFIAAYNVLSGIVITLPTLTAISFFIYFLAFLVLLFATFITQATRKIPQSSGQ--  
15mp EGPIAAYSVLSGVVVTINATLTIAISFFIYFMAFVTLFATFITQETRKIPQSSGQ--  
17mtu AEGQ-----SILE-----SRGGVFTAVCAALIIIVGVVFEVQGRRIIPVQYAKR--  
18nm SGIAKLVT-LTNQG-----SMSMLTALFIVFGALLIYLVVYFESAQRKIPHYAKR--  
19pae RAIGQSFE-SARQG-----DINIFALIGVLLAVIAIAVVFVIERGQRRIAVHYAKR--  
21rp SATISMFE-LSRKC-----ALSPLIAITVCGVLLIAIIIFFEKAQRKLLVQYPKR--  
22syn QTLGQTIIEYAQSGG-----RQSITAVVLLMLVLMIVGIVFVQEGTRRIPISARR--  
23tm SYIR-----QAYLG-----GLNLEWIFLIAVALITIFGILVQQAERRITIQYARR--  
24tp HALAEMWR-LQRLG-----ELNMVFIIVAFVFMVGIIVLVVYEQGQRRKIPHYARR--  
25vib KAIGQTIIE-QARQG-----ELHVLVLLIIVLAFVAVIYVVFVERGQRRIVVNYAKR--  
26xyf GSVLQIID-AFRSD-----ALSILSLLIIVFVLSFTFFVVFVERGLRRITVNYARR--

01aae -QVGRQLYAG-RKTYLPKINPAGVPIIIFAQALLIPSTLLNFVQ-NP-----FIKV  
04bb -NSN-----NT-VSSYLPKLNPSGVLPVIFASVLTLPQLISGFAETS-----SIARQ  
05bs --TGRSPAGGGQSTHLPKLVNPAVPIVIFAVAFILTPRTIASFFGTN--DV-TKWIQN  
06cj -VVMQNKQR-IMNYIPKLNLSGVIPPVIFASAILMFPPTILQTST-NP-----YLQA  
07cp -VIGRREVPG-GGSYLPKLVNPAVPIVIFASILLMFPATIGQFIASES-----SWMKR  
08ct -IIGRREVPG-GGSYLPKLVNPAVPIVIFASILLMFPATIGQFIASES-----SWLKR  
09der GAAGQARGAG-QATWLPKLVNPAVPIVIFASAMLIIPNLIASATATRAP-EV-NAWINS  
10ec -QQGRRVYAA-QSTHLPKLVNPAVPIVIFASSIILFPATIASWFG-GG-----TGWNWLT  
11hi -QQGRQILGG-HSTHLPKLVNMANVMPAIFASSIILFPATLTQWFGQND---K-FEWLNN  
12hp -VVMQNKQR-IMNYIPKLNLSGVIPPVIFASALLVFPSTILQQTASNK-----TLQA  
13mg ---GLVSEVK-TLPYLPKLVNPAVPIVIFASSIMSIPVTIAQFPQTES----RWFE  
15mp ---GLVTESS-ALPYLPKLVNPAVPIVIFASSIMSIPVTIAQFPQTES----RWFE  
17mtu -MVGRMYGG-TSTYLPKLVNPAVPIVIFASSLIIPHLITQLIRSGSG-VVGNSSWWDK  
18nm -QFNCR-AGS-QNTHMPFKLNPAVPIVIFASSIILFPSTLLGWGFSADT---NSVLHK  
19pae -QQGRKFVAA-QTSHLPKLVNPAVPIVIFASSILLFPASLGSWFGQSE---G-LGWLQD  
21rp -QVGNKIYGG-EATHMPKLNNTSGVIPPVIFASSILLFPPTLASFSNSNS-----DTMSM  
22syn -QVQKKLYRE-RTSYLPKLVNPAVPIVIFASVILPSSLAGFATGNEGLGFGFIVQ  
23tm -VTGRRVYGG-ASTYLPKLVNPAVPIVIFASAIIVPSAISITN-----NETLKN  
24tp -VVGKMYGG-QSTYIPKLVNPAVPIVIFASSLTFPLQIASSIGPNV-----RFLHQ  
25vib -QGRKVFVAA-QSTHLPKLVNPAVPIVIFASSIILFPGLTAQWFGQNGE-SSTFGWLD  
26xyf -VEGRNSYMN-KSSFPLKLVNPAVPIVIFASSILAFVPTLSMWGQSSANTVFGAWLQR



|       |                                                               |
|-------|---------------------------------------------------------------|
| 01aae | IADMfQ--PGAIFYNFLYVTFIVFFTYFYTAVINPVELAENLHKAGAFIPGVRPGQDTV   |
| 04bb  | ILSYLR--PNGFYTFNLVILIIIFTYFYYSKIQLSPKDISNNIRKNGGTIPGK-SDEME   |
| 05bs  | NFDN-----THPVGMAIYVALIIAFTYFYAFVQVNPQEQMADNLKKGQGYIPGVRPGKMTQ |
| 06cj  | INDFLN--PNGYLPHVLTFLFVIFFAFYFASIVFNKDAENLKKGGGIPGIRPGEQTS     |
| 07cp  | IAALLA--PGSLVYSICYVLLIIFFTYFYWTATQFHPQEQIASEMKKNNAFIPGIRQKPTQ |
| 08ct  | IATMLS--PGSVAYSIFYVLLIIFFTYFYWTATQFRPEQIASEMKKNNAFIPGIRQKPTQ  |
| 09der | NLTF-----CQPLYLALESLLIFGFTYLYNSVQFDPKRIAEQLREAGGFIPGVRPGGPTA  |
| 10ec  | ISLYLQ--PGQPLYVLLYASAIIFCFFYTALVFNPRETADNLKKSAGFVPGIRPGEQTA   |
| 11hi  | LSMLLN--PGQPLYLLVYAVAIIFFSFFYTAMQYNPRDTADNLKKSAGFIPGIRPGEQTS  |
| 12hp  | VADFLS--PQGYAYNILMFLIIFFAYFYSSIVFNKDIADNLRRNGGYIPGLRPGEGTS    |
| 13mg  | DYLSLS----TPVGIFLYAVLVILFSFFYSYIQINPERLAKNFEKSGRFIPGIRPGNDTE  |
| 15mp  | DYLSLS----KPTGIVLYGILVILFSFFYSYIQINPERLAKNFEKSGRFIPGIRPGKDE   |
| 17mtu | FVGTLYSDPSNLVYIGIYFGLIIFFTYFYVSITFNPDERADEMKFGGFIPIGRGRPTA    |
| 18nm  | IAGLLQ--HGQLLYMALFAATVIFFCYFYTALVFSPEKMAENLKKSGAFVPGIRPGEQTS  |
| 19pae | VQAATA--PGQPLNILLFTAGIVFFCFFYTALMFNPKDVAENLKKSGAFIPGIRPGEQSA  |
| 21rp  | LTYYLG--HGKPYVILLVYVLMFSSFFYTAVFNSEETANNLRKYGAYIPGRKPKNTS     |
| 22syn | ISNALR--PCTWVYTVVYSVMIFFFSYFYASLVNPEVSNLKNMGSSIPGIRPGKTE      |
| 23tm  | LFRA-----GGFLYLLIYGLLVFFTYFYVSIVFDPREISENIRKYGGYIPGLRPGRSTE   |
| 24tp  | LAQFLR--PNSWYNAFYVVLIVFFAYFYTQVTLNPTETAKQIRENGGTIPGIR-ADKTE   |
| 25vib | VSLALS--PGQPLYVMLYAAAAIIFCFFYTALVFNPRETADNLKKSAGFVPGIRPGEQTA  |
| 26xyf | VSSALG--PGEPLHMLFAVLIIGFAFFYTALVFSSQETADILKKSAGALIPGIRPGKATS  |

|       |                                                               |
|-------|---------------------------------------------------------------|
| 01aae | KYLERIINRLIFFGALFLSVIALIP--ILISVWFNIP---FYFGGTTALIVVGVALDTF   |
| 04bb  | KYLDEIMNKTFLFSGSIFLSIIAIIIP--FLVQNIFRPHDVSRIIMGSSLLIMVGVALDTL |
| 05bs  | DRITSIIYRLTFVGSIFLAVISILP--IFFIQFAGLPQS--AQIGGTSLLIVVGVALETM  |
| 06cj  | SYLNEVASRLTLGSGIYLGIVATLP--WVLVKFMGVP----FHFGGTSVLIVVQVALDTM  |
| 07cp  | HLEYTMNRVTLLGALFLAAIAIIP--SLLGCLLRVDSNVSYFLGGTAMLIIVGVVLDTM   |
| 08ct  | TYLEYTMNRVTLLGAVFLAVVAILP--SILGRILRVDANVSYFLGGTAMLIIVGVILDTM  |
| 09der | DFLGRISGRSLWGAIIVILITILP--QLVQKATGITT---FQFSGTGLLIIVGVALETL   |
| 10ec  | KYIDKVMTRLTLVGALYITFICLIP--EFMRDAMKVP---FYFGGTSLLIVVVVIMDFM   |
| 11hi  | RYIDKVMTRLTLIGGLYVTFVCLVP--YIMTSAWDVK---FYFGGTSLLIVVVVIMDFI   |
| 12hp  | SFLNSVASKLTLWGSYLALISTVP--WILVKAMGVP----FYFGGTAVLIVVQVAIDTM   |
| 13mg  | KHIAARVLIRINFIGAPFLTVIAIIP--YIVSYFIRLPN--SLSLGGTGIIIVTAVVEFI  |
| 15mp  | KHIAARVLIRINFIGAPFLTVIAIIP--YIVSALIHLPN--SLSLGGTGIIIVTAVVEFM  |
| 17mtu | DYLRVLSRLTLPGSIYLGIVAVLPNLFQIAGGTVQN--LPFGGTAVLIMIGVGLDTV     |
| 18nm  | RYLEKVVLRLLTFLGALYITFICLIP--EFLTTLVNLVP---FYLGGTSLLIVVVVIMDFS |
| 19pae | RYIDGVLTRLTMFGALYMTAVCLLP--QLFVVAHVVP---FYLGGTSLLIVVVVIMDFM   |
| 21rp  | DYFDYILTRLTVIGGIYLSVICVIP--ELLMNKYVIS---LSLGGTSFLIVNVVLDTM    |
| 22syn | QYLEGVNLRLTLGALFLSFVATLP--IFVEQATGVTT--FQGLGATSLILVGVADTA     |
| 23tm  | QYLHRVLNRVTFIGAVFLVVIALLP--YLVQGAIKVN---VWIGGTSALIVGVADLII    |
| 24tp  | EYLGQILNRLVLPGLSYLGMIAVLP--TLIQAAGFPSSISLLMGGTSLLILVGVLDLTM   |
| 25vib | KYIDKVMTRLTLAGALYITFICLIP--EFMMVAWNVN---FYFGGTSLLIVVVVIMDFM   |
| 26xyf | DYVDGVLTRLTAVGSLYLIVICLIP--EFMQSRMGAS---FHFGGTSLLIAVVVIMDFI   |

|       |                                                 |
|-------|-------------------------------------------------|
| 01aae | RQIETYLIQKKYKSY-----VRR-----                    |
| 04bb  | IHIDAYLKT---QGFSH-----GNKKN---YAFLLQKI-----     |
| 05bs  | KQLESQVLKRNRYRGF-----MKN-----                   |
| 06cj  | RKIEAQIYMSKYQTL-----SAVGL-----                  |
| 07cp  | KQVDAFLLMRRYDSV-----LKTDRTKGRH-----             |
| 08ct  | KQIDAFLLVRRYDGV-----LKKDRPKGRP-----             |
| 09der | KQLEAQLTVRRYDGF-----ISKGRLRGRLLN-----           |
| 10ec  | AQVQTLMMSSQYESA-----LKKANLKGYGR-----            |
| 11hi  | VQVQSHLMSSQYESA-----LKKANLKGFGQ-----            |
| 12hp  | KKIEAQIYMSKYKTL-----SAVGF-----                  |
| 13mg  | SALRSAATATNYQQLRRNLAIEVQQTAKQDSLEQLQKEAPGIGNLW  |
| 15mp  | SALRSAATATNYQQLRRNLAIEVQQTAKQDKEEQLEAETPGIGNLW  |
| 17mtu | KQIESQLMQRNYEGF-----LK-----                     |
| 18nm  | TQINSYRLTQQYDKL-----MTRSEMKSFSRK-----           |
| 19pae | AQVQSHLVSHQYESL-----MKKANLKGYSGMLR-----         |
| 21rp  | TQIQTYLFSKYEGL-----MKKIKLKN-----                |
| 22syn | KQIQTYLVISQRYEGM-----IKQP-----                  |
| 23tm  | QQMETHMVMRHYEGF-----IKKGKIRGR-----              |
| 24tp  | SQIEAQLKMRQREGLGG-----RGKVLPRICSGYLKRDVYEE----- |
| 25vib | AQVQTHLMSHQYESV-----LKKANLKGYGR-----            |
| 26xyf | SQIQVHQASNQYDGVGF-----LKKANLKGVSRRRRFVN-----    |

# Anhang D

## Der Plastidendatensatz

01can MSNIRPDEISSILKQIERYNESVKIENTGTVLQVGDGIARIYGLDNIMAGELLEFEEDKT  
02chl MVKIRPDEISSIIKQIEQYQEVKAVNVGTVFQVGDGIARIYGLDKVMAGELVEFEDGT  
03cya MVSIRPDEISSIIRQQIEQYDQEIQVSNVGTVLQVGDGIARVYGLDKVMSGELLEFEEDGT  
04eug MIRVRPNEVTRIIRQQVKKYRQELKIVNVGTVLQVGDGIARIYGLEKVMAGELVEFEDGT  
05gui MVNIRPDEISSIIRQQIDKYDQAIQVSNVGTVLQVGDGIARVYGLDQVMAGELLEFEEDKT  
06mai MATLRVDEINKILRERIEQYNRKVGIEIGRNVVQVGDGIARIYGLGEIMSGELVEFAEGT  
07mar MVNIRPDEISSIIRKQIEQYNQEVKIVNIGTVLQVGDGIARIYGLDKVMAGELVEFEDGT  
08mes MIKIQPEEISSVIRKQIEQYNQEVKVVNTGTVLQVGDGIARIYGLAKAMAGELLEFEEDGT  
09nep MVKIRPDEISNIIIRQQIEQYSQEVKVVSVGTVLQVGDGIARIYGLEKVMAGELLEFEEDGT  
10nic MVTIRADEISNIIIRERIEQYNREVKIVNTGTVLQVGDGIARIHGLDEVMAGELVEFEEDGT  
11odo MINIRPDEISSIIRQIEQYDQVKVDNIGTVLQVGDGIARVYGLDQVMSGELLEFEEDKT  
12oen MATIRADEISNIIIRERIEQYNREVKIVNTGTVLQVGDGIARIYGLDEVMAGELVEFEEDGT  
13ory MATLRVDEIHKILRERIEQYNRKVGIEIGRNVVQVGDGIARIYGLGEIMSGELVEFAEGT  
14pin MGSIRLDEISSIIRKQIEQYNNEVRVGNLGTVLQVGDGIARIHGLDEVMAGELVEFGDGT  
15por MVNIRPDEISSIIRQQIEKYDQDEVANIGTVLQVGDGIARVYGLDEVMAGELLEFEEDKT  
16syn MVSIRPDEISSIIRQQIESYDQSVQVSNVGTVLQVGDGTARIYGLEQVMSQELLEFEEDGT

01can IGIALNLETDNVGAVLMGDRDILEGSSVKGTGKIAQIGVGNLLGRVNLALGNPIDGKP  
02chl VGIALNLEAKNVGAVLMGEGTRVQEGSSVRATGKIAQIPVGDGYLGRVNSLARPIDGKG  
03cya IGIALNLEADNVGVVLMGDRNLLEGSSVRATGKIAQIPVGDGAVIGRVVDALARPIDGKG  
04eug IGIALNLEADNVGAVLMGEATNLKEGASVKTGKIAQIPVGRGFLGRVVDALARPIDGKG  
05gui IGIALNLESDNVGVVLMGEGRGILEGSSVKATGKIAQIPVPGKSYLGRVINALGTPIDGKG  
06mai RGIALNLESKNVGIIVLMGDLMIQEGSFVKATGRIAQIPVSEAYLGRVINALAKPIDGRG  
07mar VGIALNLESDNVGAVLMGDGLTIQEGSSVKATGKIAQIPVSDAYLGRVINALAKPIDGKG  
08mes VGIALNLESNVNGAVLMGDGFSIQEGSRVKATGKIAQIPGESYIGRVVDALARPIDGKG  
09nep VGIALNLEADNVGAVLMGSGLSIQEGSAVKATGKIAQIPVGEAFLGRVINALARPIDGKG  
10nic IGIALNLESNVNGVVLMDGLLIQEGSSVKATGRIAQIPVSEAYLGRVINALAKPIDGRG  
11odo IGIALNLENDNVGVVLMGNRGRIQEGSTVKTGTGIAQIPTGEAFLGRVNVPLGAPIDGKG  
12oen IGIALNLESKNVGVVLMGDGLLIQEGSSVKATGRIAQIPVSEAYLGRVINALAKPIDGRG  
13ory RGIALNLESKNVGIIVLMGDGLMIQEGSFVKATGRIAQIPVSEAYLGRVINALAKPIDGRG  
14pin VGIALNLGSDNVGAVLMGDGLMIQEGSSVRATGKIAQIPVSDAYLGRVINALAKPIDGKG  
15por IGVALNLESDNVGVVLMGDRDILEGSSVKGTGKIAQIPVGD AFLGRVVDPLARPIDNKG  
16syn IGIALNLEEDNVGAVLMGDGFGIQEGSTVKTGTGIAQIPIDGAMVGRVVDLGRPIDGKG

01can NPNSTDYRLIEFNAPGIVSRRSVCEPIQTGITAIADAMIPIGRGQRELIIGDRQTGKTSIA  
02chl EIATKENRLIESPAGIISRRSVHEPLQTGIVAIDAMIPIGRGQRELIIGDRQTGKTAIA  
03cya DIATTDTRLIESAPGIIISRFVYEPLQTGITAIADAMIPIPRGQRELIIGDRQTGKTAVA  
04eug DIASFTRRLIESPAGIISRRSVHEPLQTGLIAIDAMIPIGRGQRELIIGDRQTGKTAVA  
05gui DINCSETRLIESIAPGIIISRSVCEPIQTGITAIIDSMIPIGRGQRELIIGDRQTGKSSVA  
06mai EIVASESRLIESPAGIISRRSVYEPLQTGLIAIDSMIPIGRGQRELIIGDRQTGKTAVA  
07mar QIPASEFRLIESPAGIISRRSVYEPMTGLIAIDSMIPIGRGQRELIIGDRQTGKTAVA  
08mes DIPSSSETRLIESPAGIISRRSVYEPLQTGLVSVSDAMIPIGRGQRELIIGDRQTGKTAVA  
09nep DIASKESRLLEGAPGIIERRSVYEPMTGLIAIDAMIPIGRGQRELIIGDRQTGKTAIA  
10nic EISASEFRLIESAAPGIIISRRSVYEPLQTGLIAIDSMIPIGRGQRELIIGDRQTGKTAVA  
11odo DIANTETRLLEAMAPGIIISRSVCEPLQTGITSIDAMIPIGRGQRELIIGDRQTGKTAIA  
12oen EISSSESRLIESPAGIISRRSVYEPLQTGLIAIDAMIPIGRGQRELIIGDRQTGKTAVA  
13ory EIVASESRLIESPAGIISRRSVYEPLQTGLIAIDSMIPIGRGQRELIIGDRQTGKTAVA  
14pin KISASEFRLIESPAGIISRRSVYEPLQTGLIAIDSMIPIGRGQRELIIGDRQTGKTAVA  
15por EPASNGTRLIESMAPGIIIGRSVCEPMQTGITAIIDSMIPIGRGQRELIIGDRQTGKTAVA  
16syn PISSTATRLLESAPGIIIRKSVCEPMQTGITAIADAMIPIGRGQRELIIGDRKTGKTAIA

01can LDITIINQKEENVICIVYVAIGQKASSVAQAVTLLEEKDALKYTVVIAANANEPATLQYIAP  
02chl VDTILNQKQGDVVCVYVAIGQKASSIAQVNVNLTQERGAMDYTIIVAANADSPATLQYLS  
03cya IDTILNQKQGDVVCVYVAIGQKASSVAQVGVLEKQALDYTVIVAANADDPATLQYLAP  
04eug TDTILNQKQGDVVCVYVAIGQKASSVSQIVTTLLEKRGAMEYTIIVAENADSSATLQYLAP  
05gui IDITIINQKGEDVVCVYVAVGQKAATVASIVTTLLEEKALDYTCIVAANADDPATLQYIAP  
06mai TDTILNQKQGDVVCVYVAIGQRASSVAQVVTTFHEEGAMEYTIIVAENADSPATLQYLAP  
07mar IDTILNQKQGDVVCVYVAIGQKASSVAQVNVNLTQERGALDYTVIVAANADSPATLQYLAP  
08mes IDTILNQKQGDVVCVYVAIGQKASSVAQVVTSTLEENGAMAYTIIVAENANAPATLQYLAP  
09nep TDAIVNQKQSGVICVYVAIGQKASSVAQIVTTLQEKDAMRYTIIVSETADSPATLQYLAP  
10nic TDTILNQKQGNVICVYVAIGQKASSVAQVVTTLQERGAMEYTIIVAETADSPATLQYLAP  
11odo VDTIINQKTEDVVCVYVGQKASTVAQVNVNLEEKAMAYTIIVCASANDPATLQYIAP  
12oen TDTILNQKQGNVICVYVAIGQKASSVAQVNVNLTQERGAMEYTIIVAEEADSPATLQYLAP  
13ory TDTILNQKQGDVVCVYVAIGQRASSVAQVVTTFHEEGAMEYTIIVAENADSPATLQYLAP  
14pin TDTILNQKQSNVICVYVAIGQRASSVAQVNVNLTFRERGAMAYTIIVAETADSPATLQYLAP  
15por LDITIINQKQGDVVCVYVAIGQKASSVAQVSSLEKQALDYTIIVTANADSPATLQYIAP  
16syn IDITIINQKQSEDVICVYVAIGQKASTVAQIIDTLTEKGAMAYTIIVAANADDPATLQYLAP



|       |                                                               |
|-------|---------------------------------------------------------------|
| 01can | YTGAATAEHFMYKGLATLIIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE    |
| 02chl | YTGAALAEYFMYTGRHTLVIYDDLTKQAQAYREMSLLLRPPGREAYPGDVFLHSRLLE    |
| 03cya | YTGASIAEYFMYKGHTLVIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE     |
| 04eug | YTGAALAEYFMYNGKHTLVIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE    |
| 05gui | YTGAALAEYFMYNGATLVIYDDLTKQAQAYREMSLLLRPPGREAYPGDVFLHSRLLE     |
| 06mai | YTGAALAEYFMYRHTLIIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE      |
| 07mar | YTGAALAEYFMYRKHTLIIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE     |
| 08mes | YTGATLAEFFMYSGRHTLVIYDDLTKQAQAYREMSLLLRPPGREAYPGDVFLHSRLLE    |
| 09nep | YTGAALAEYFMYSGRHTLVIYDDLTKQAQAYREMSLLLRPPGREAYPGDVFLHSRLLE    |
| 10nic | YTGAALAEYFMYRHTLIIYDDPSKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE      |
| 11odo | YAGAALAEYFMYNGATLVIYDDLTKQAMAYRQMSLLLRPPGREAYPGDVFLHSRLLE     |
| 12oen | YTGAALAEYFMYRHTLIIYDDPSKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE      |
| 13ory | YTGAALAEYFMYRHTLIIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE      |
| 14pin | YTGATLAEYFMYKKQHTSIIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE    |
| 15por | YTGAALAEYFMYKGATLVIYDDLTKQAQAYRQMSLLLRPPGREAYPGDVFLHSRLLE     |
| 16syn | YTGATLAEHFMYQKSTLVIYDDLTKQAQAYRQMSLLMRPPGREAYPGDVFIHSRLLE     |
|       |                                                               |
| 01can | RAAKLNNEGGGSM TALPIIETQAGDVSAYIPTNVISITDGGIFLSSDLFNAGIRPSINV  |
| 02chl | RAAKLNDKLGSGSM TALPVVETQEGDVSAYIPTNVISITDGGIFLSADIFNAGIRPAINV |
| 03cya | RAAKLSPQLGEGSM TALPIVETQAGDVCAYIPTNVISITDGGIFLSADLFNSGLRPAINV |
| 04eug | RAAKLSDGLGEGSM TALPIVETQAGDVSAYIPTNVISITDGGIFLSADIFNSGIRPAINV |
| 05gui | RAAKLSDKLGGGSM TALPVVETQAGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
| 06mai | RAAKLNSLLGEGSM TALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
| 07mar | RAAKLSSNLGEGSM TALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
| 08mes | RAAKLSNALGEGSM TALPIIETQAGDVAAYIPTNVISITDGGIFLSADLFNSGIRPAINV |
| 09nep | RAAKLSDALGEGSM TALPVVETQGGDVSAYIPTNVISITDGGIFLSADIFNAGIRPAINV |
| 10nic | RAAKLSSSLGEGSM TALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNSGIRPAINV |
| 11odo | RAAKLSDALGGGSM TALPVVETQASDVSAYIPTNVISITDGGIFLSADLFNSGIRPAINV |
| 12oen | RAAKLSSRLGEGSM TALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
| 13ory | RAAKLNSLLGEGSM TALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
| 14pin | RAAKLSSQLGEGSM TALPIVETQAGDVSAYIPTNAISITDGGIFSSADLFNAGIRPAINV |
| 15por | RAAKLNSDLGGGSM TALPIIETQAGDVSAYIPTNVISITDGGIFLSADLFNSGIRPAINV |
| 16syn | RAAKLSDALGGGSM TALPVVETQAGDVSAYIPTNVISITDGGIFLSADLFNAGIRPAINV |
|       |                                                               |
| 01can | GISVSRVGSAAQIKAMKQVAGKLELAQFDELQAFSDFASDLDKSTQAQLARGQRLREL    |
| 02chl | GISVSRVGSAAQPKAMKQVAGKLELAQFAELAFSDFASDLQDQATQNQLARGQRLREL    |
| 03cya | GISVSRVGSAAQIKAMKQVAGKLELAQFDELKAFSDFASDLDKATQQLARGRLREL      |
| 04eug | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLQDQATQNQLARGRLREL     |
| 05gui | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLQDQATRNQLARGQRLREI    |
| 06mai | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKTSQNQLARGRLREL      |
| 07mar | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQNQLARGQRLREL     |
| 08mes | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQNQLARGRLREL      |
| 09nep | GISVSRVGSAAQVAMKQVASKLELAQFSELEAFSDFASDLDAATQAQLARGVRLREL     |
| 10nic | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQNQLARGQRLREL     |
| 11odo | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDEATQQLARGTRLREV      |
| 12oen | GISGSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQNQLARGQRLREL     |
| 13ory | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKTSQNQLARGRLREL      |
| 14pin | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQDQLARGQRLREL     |
| 15por | GISVSRVGSAAQIKAMKQVAGKLELAQFAELAFSDFASDLDKATQNQLARGQRLREI     |
| 16syn | GISVSRVGSAAQTKAMKQVAGKLELAQFAELAFSDFASDLDAATQAQLARGQRLRQL     |
|       |                                                               |
| 01can | LKQPQASPISVVEEQIPIYAGINGFLDDIKIDRISLFIKKLQECLSNSYPHFYEAIKESK  |
| 02chl | LKQSSQSPLSLEDQVASIYAGTNGYLDVLPADRVRAFLVGLRQYLATNKAKEYGEILRSTN |
| 03cya | LKQQQYAPLPVEEQVAVIYTGINGFLDNIEQKVSFAFISNLRNLSVKRAKFGEIRSEK    |
| 04eug | LKQPQASPLSVADQVATIIYTGINGYLDINLEDVGRGLIELREHINNEKPTFKEIINKTK  |
| 05gui | LKQPQNSPISVEEQVAIIYTGINGYLDIADVKKRRVFTNLRNLSKNSKPQYAEIRNTK    |
| 06mai | LKQSSQNSPLPVEEQVATIIYTGTRGYLDSLEIEQVKKFLDELKHLKDTKPFQFEIISSSK |
| 07mar | LKQSSQAPLSVEEQIATIIYTGNGYLDVLETGQVKKFLIQLREYVLTNKPQFAEIRSTK   |
| 08mes | LKQAQSSPLPVAQVLTIIYAGVNGYLDIAIEDVKKFLAGLRNVIITSKAAIINNINSSK   |
| 09nep | LKQAQSEPLSVADQVATIIYTGNGYLDLAPTQVRAFLSALRSYLATSKPKYAQIM-AAN   |
| 10nic | LKQSSQAPLTVEEQIMTIYTGNGYLDLSLEVGGVKKFLVELRVTYKTNKPQFEIISSTK   |
| 11odo | LKQPQNSPLSVAEQVALIYTGINGFLDELEVASVKKYCASLSFLNLSNNSYIGIVSSTN   |
| 12oen | LKQSQAKPLTVAEQILTIYTGNGYLDSEFAIQVRKFLDELDRYVVKTRKPFQFEIISSTK  |
| 13ory | LKQSQANPLPVEEQIATIIYTGTRGYLDSLEIGQVKKFLDELKHLKDTKPFQFEIISSSK  |
| 14pin | LKQSSQAPLTVEEQIATIIYTGNGYLDIFEIAQVRKFLGLRFLYLIKNNKPQFEIRSTG   |
| 15por | LKQAQNSPISVEEQVAIIYTGINGYLDIADVKNVPDFIKLREDLKNKSPFEGEIRSSK    |
| 16syn | LKQPENSPLSVWEQVAISYAGLNGYIDTIPVDKVFTEFAQLRDYLNKANKAKYVEINSSK  |
|       |                                                               |
| 01can | QLSKENEVLLKKAITEVKKNLSFI---MKPCQLKQKQKTLAIKVNKNKGVIQIIGPVLD   |
| 02chl | ALTDEAQTLLKEALKEYTEEFASAK-----MSVS-TETKSSGRITQIIGPVLD         |
| 03cya | ALTAEAEENLLKDAISDCKQAFLSNI-----MATTTSKTNTGYVTQVIGPVLD         |
| 04eug | TFTEAEALLKFTIIDLKKNFKKRIK-----MKSSRLRLNTGIIILQIIGPVMD         |
| 05gui | TFNSDAENLLKSAIADTKQSFV-----MTNTATNTIGYITQIIGPVVD              |
| 06mai | TFTEQAETLLKEAIEQLERFSLQEQT-MRTNPPTSRPGISTIEEKSVRIDQIIGPVLD    |
| 07mar | VFTEQAENLLKEAITEHIELFLFQEEK-MKTNFLAFG--MSTLVAKNIGSITQVIGPVLD  |
| 08mes | AVTPETEGMMKEAINEYKVFVFAAG-----MTAIEETRIGSITQIIGPVLD           |
| 09nep | VFTPEAESLVKEAIAETKASFK-----MSQAGVKTKQAGIKVQIGPVMD             |
| 10nic | TFTEEAALLKKEAIEQMDRFLQEQA-MRINPTTSGSGVSTLEKKNPGRVVQIIGPVLD    |
| 11odo | QFTAEAEATLKKEAISESKAVFMK-----MVKTNINKGVNQIIGPVLD              |
| 12oen | IFTTEEAQALLKDAIHEQKELFLVQE---MRINPTTSGPGVSTLEKKSGRIAQIIGPVLD  |
| 13ory | TFTEEAELLKKEAIEQLERFSLQEQT-MRTNPPTSRPGVSTIEEKSRTGRIQIIGPVLD   |
| 14pin | TFTEEAKALLEAL-----MRKNPLVLG--VSASVETNVGRIAQIIGPVLD            |
| 15por | KLDTASEELLKKAIEDVKQGFVKA-----MVSTTKSTGSVTQIIGPVLD             |
| 16syn | ALTDEAETLLKEGIEFTQGFAG-----MVAVKEATNVGKITQVIGPVLD             |

|       |                                                                |
|-------|----------------------------------------------------------------|
| 01can | IVFPDGLPKVFNAINNSNNN----WITCEVQQLLDGNKVRVAVAMSTTEGLKRGASAI     |
| 02chl | VSPFAGKMPNIYNALTVGGKNEAGQEISVTCEVQQLLDGHCVRVAVAMSATDGLMRGMEVL  |
| 03cya | VSPFNGQLPKIYNALTVKGKNEAGQDITVTCEVQQLLDGNQVRVAVAMSTTDLGRGMEVT   |
| 04eug | ISFSPGKMPNIYNALLIEGKTESGDRKVVCEVQQLLDGNVVRAIAMSATDGLRGIKIVI    |
| 05gui | VEFTTGKLPQIYNALIIIGSGEN----AVTCEVQQLLDGNKVRVAVAMSTTDLGRGMEVT   |
| 06mai | ITFPPGKLPYIYNALIVKSRDTADKQINVTCEVQQLLDGNRVRVAVAMSATDGLMRGMEVI  |
| 07mar | VAFSPGKMPNIYNALIVKDQNSAGEEINVTCEVQQLLDGNKVRVAVAMSATDGLMRGMEVI  |
| 08mes | AKFPPGQMPNIYNALIVKAKNEAGEDLSVTCEVQQLLDGNQVRVAVAMSATDGLMRGLEIL  |
| 09nep | VAFPPGSMMPNIYNALIVKGRNGAGQEVSVCEVQQLLDGDLVRVAVAMSATDGLMRGMEVT  |
| 10nic | VAFPPGKMPNIYNALVVQRDSVGGQINVAECVQQLLDGNRVRVRAIAMSATDGLTRGMEVI  |
| 11odo | IEFPPSGTLPPISALKIETEDG---IGTVVEVQQLLDGNKVRVAVAMSTTDLKRGVEAR    |
| 12oen | VTFFPGKMPNIYNALVVKGRDTGGQEIINVTCEVQQLLDGNRVRVAVAMSATDGLTRGMEVI |
| 13ory | VTFFPGKLPYIYNALVVKSRDTDGKQINVTCEVQQLLDGNRVRVAVAMSATDGLMRGMEVI  |
| 14pin | VSPFPGNMPNIYNALIVKGQGTAGQEIVTCEVQQLLDGNKVRVAVAMSATDGLTRGMRVI   |
| 15por | IAFPNGQLPKVFNALKVQSTDG----TITCEVQQLLDGNKVRVAVAMSTTEGLKRGVEVV   |
| 16syn | AQFPSPGKLPRIYNALVKQRNSAGNEVAVTCEVQQLLDGNQVRVAVAMSTTDLVRGMDVV   |
|       |                                                                |
| 01can | DTGEPISIPVGKETLGRIFNVLGEPIDEGPVISNDKLPiHRPAKFTQLETKPSIFETG     |
| 02chl | DSGKPLNVPVGAATLGRIFNVLGEPVDNLGPVNTKQLPIHRSAPAFVDLTKLSIFESG     |
| 03cya | DSGAAISVPVGTPTLGRIFNVLGEPVDELGAVCDSTLPIHRSPSAFTQLETKSSIFETG    |
| 04eug | DTGAPLSVPVGGTTLGRIFNVLGESVDKMGIDYSITLPIHRAAPSFVELDTKLSIFETG    |
| 05gui | DTGAPISVPVGGATLGRIFNVLGEPVDNMGDVALDQTLPIHRGSAFTQLETKPSIFETG    |
| 06mai | DTGTPLSVPVGGATLGRIFNVLGEPIDNLGPVDTSATFPIHRSAPAFIELDTKLSIFETG   |
| 07mar | DTGAPLTPVPGEATLGRIFNVLGEPVDNLGPVEVTTTPIHRAAPAFQLDTKLSIFETG     |
| 08mes | DTGAPLSVPVGEITLGRIFNVLGEPVDGLGDVAKTKSPIHRNSPTFIELDTKLSIFETG    |
| 09nep | DTGRALSVPPVGGTTLGRIFNVLGEPVDNMGPVGNEKTLPIHREAPAFVDLTKLSIFETG   |
| 10nic | DTGAPISVPVGGATLGRIFNVLGEPVDNLGPVDTSTTSPHRSAPAFIELDTKLSIFETG    |
| 11odo | DLGGPISVPVGI STLGRIFNVIGEPVDEQGDVSYDETLPiHREAPAFTELETKPSIFETG  |
| 12oen | DTGAPLSVPVGEATLGRIFNVLGELVDELGPVDTTRTSPHRSAPAFIELDIELSIFETG    |
| 13ory | DTGAPLSVPVGGATLGRIFNVLGEPVDNLGPVDTSATFPIHRSAPAFIELDTKLSIFETG   |
| 14pin | DTGAPLSVPVGGATLGRIFNVLGEPVDNLGPVDAGITSPiHRPAAPAFELDTKLSIFETG   |
| 15por | DTGAPISVPVGTNTLGRIFNVLGEPVDNLGPVSSESTLPIHRPAAPFTKLETKPSIFETG   |
| 16syn | DTGAPISVPVGTGTLGRIFNVLGEPVDNKGPPVAGETFPiHRPAKLVDELTKPQVFETG    |
|       |                                                                |
| 01can | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNVAKAHGGVSVFGVGERTREGNDL    |
| 02chl | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 03cya | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 04eug | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 05gui | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 06mai | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 07mar | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 08mes | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 09nep | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 10nic | IEVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 11odo | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 12oen | IKVVDLLAPNRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 13ory | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 14pin | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 15por | IKVVDLLAPYRRGGKIGLFGGAGVGKTVLIMELINNIAKAHGGVSVFGVGERTREGNDL    |
| 16syn | IKVIDLLTPYRQGGKIGLFGGAGVGKTVIMMELINNIAIQHGGVSVFGVGERTREGNDL    |
|       |                                                                |
| 01can | YQEMKESGVINEKDNLNSKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQVLLFI    |
| 02chl | YMEMKESGVINESNISKVALVYGGQMNPPGARMRVGLTALTMAEYFRDINKQDVLLFI     |
| 03cya | YQEMKESKVIDENWLPASKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNNQDVLLFI   |
| 04eug | YQEMKESGVINDRNFKESKVALIYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI   |
| 05gui | YMEMKESKVINESNLGESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI   |
| 06mai | YMEMKESGVINEKNIESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI    |
| 07mar | YMEMKESKVINEQNIIESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI   |
| 08mes | YMEMKESKVINEENLSSKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNNQDVLLFI    |
| 09nep | YMEMCESKVINKADVKSVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNNQDVLLFI     |
| 10nic | YMEMKESGVINEENIAESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNEQDVLLFI   |
| 11odo | YEEMKESGVINEENNFESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI   |
| 12oen | YMEMKESGVINEQNIIESKVALVYGGQMNPPGARMRVGLTAVRVADYFRDVNKQVLLFI    |
| 13ory | YMEMKESGVINEKNIESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI    |
| 14pin | YMEMKESGVIDEQKIESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDVNEQDVLLFI    |
| 15por | YMEMKESKVINEDNLKESKVALVYGGQMNPPGARMRVGLTALTMAEYFRDINKQDVLLFI   |
| 16syn | YNEMIESNVINADKPEESKIALVYGGQMNPPGARMRVGLTALTMAEYFRDVNKQDVLLFI   |
|       |                                                                |
| 01can | DNIFRFVQAGSEVSALLGRMPSAVGYPQLGTEMGALQERITSTLDGSITSIQAVVVPAD    |
| 02chl | DNIFRFVQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTKDGISTSIQAVVVPAD     |
| 03cya | DNIFRFVQAGSEVSALLGRMPSAVGYPQLGTEMGLQERITSTTKGSITSIQAVVVPAD     |
| 04eug | DNIFRFVQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTKVGISTSIQAVVVPAD     |
| 05gui | DNIFRFVQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTTEGSITSIQAVVVPAD     |
| 06mai | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKKGISTSIQAVVVPAD     |
| 07mar | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKEGSITSIQAVVVPAD     |
| 08mes | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSREMGLQERITSTKQGSITSIQAVVVPAD     |
| 09nep | DNIFRFVQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTKDGISTSIQAVVVPAD     |
| 10nic | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKEGSITSIQAVVVPAD     |
| 11odo | DNIFRFTQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTTGSITSIQAVVVPAD      |
| 12oen | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKAGISTSIQAVVVPAD     |
| 13ory | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKKGISTSIQAVVVPAD     |
| 14pin | DNIFRFVQAGSEVSALLGRMPSAVGYPQLSTEMGLQERITSTKKGISTSIQAVVVPAD     |
| 15por | DNIFRFVQAGSEVSALLGRMPSAVGYPQLATEMGLQERITSTTEGSITSIQAVVVPAD     |
| 16syn | DNIFRFVQAGSEVSALLGRMPSAVGYPQLGTVDGLQERITSTKEGSITSIQAVVVPAD     |

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01can | DLTDPAPATTFAHLDAATTVLSRLAAGKIYPVAVDPLDSTSTMLQPGIVSDEHYTTARKVK   |
| 02chl | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPWIVGDQHYQCAQNVK  |
| 03cya | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPGIVGEEKHYACAQRVK |
| 04eug | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPWIVGEEHYNTAQSVK  |
| 05gui | DLTDPAPATTFSHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQINIVGQEHYSCAQNPK  |
| 06mai | DLTDPAPATTFAHLDAATTVLSRGLASKGIYPVAVDPLDSTSTMLQPRIVGNEHYETAQRVK  |
| 07mar | DLTDPAPATTFAHLDAATTVLSRGLAAGKIYPVAVDPLDSTSTMLQPWIVGEEHYETAQGVK  |
| 08mes | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPWIVGEEHYETAQRVK  |
| 09nep | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPWILGDDHYGTAQRVK  |
| 10nic | DLTDPAPATTFAHLDAATTVLSRGLAAGKIYPVAVDPLDSTSTMLQPRIVGEEHYETAQRVK  |
| 11odo | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPGIVSGEHEYIAETVK  |
| 12oen | DLTDPAPATTFAHLDAATTVLSRGLAAGKIYPVAVDPLDSTSTMLQPRIVGDHEHYETAQRVK |
| 13ory | DLTDPAPATTFAHLDAATTVLSRGLASKGIYPVAVDPLDSTSTMLQPRIVGNEHYETAQRVK  |
| 14pin | DLTDPAPATTFAHSDAATTVLSRGLAAGKIYPVAVDPLDSTSTMLQPWIVGEEHYETAQGVK  |
| 15por | DLTDPAPATTFAHLDAATTVLSRNLASKGIYPVAVDPLDSTSTMLQPGIVGTDHYSTAQEVK  |
| 16syn | DLTDPAPATTFAHLDTTVLSRGLAAGKIYPVAVDPLDSTSTMLQPSIVSGEHYDTAREVQ    |
|       |                                                                 |
| 01can | ETLQRYKELQDIIAILGLDELSEEDRLIVSRARKIERFLSQPFFVAEVFTGISGKYVSL     |
| 02chl | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVSLA    |
| 03cya | GILQRYKELQDIIISILGLDELSEEDRLAVARARKVERFLSQPFFVAEVFTGSPGKYVSL    |
| 04eug | KTLQRYKELQDIIAILGLDELSEEDRLVSRARKVERFLSQPFFVAEVFTGSPGKYVTLA     |
| 05gui | ETLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVSLA    |
| 06mai | ETLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVGLA    |
| 07mar | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVSLR    |
| 08mes | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVSLA    |
| 09nep | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGAPGKYVPLA    |
| 10nic | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVGLA    |
| 11odo | ETLQRYKELQDIIAILGLDELSEEDRLIVARARKVERFLSQPFFVAEFTGSPGKYVSL      |
| 12oen | ETLQRYKELQDIIISILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVGLA   |
| 13ory | QTLQRYKELQDIIAILGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVGLA    |
| 14pin | QTLQRYKELQDIIAIPGLDELSEEDRLIVARARKIERFLSQPFFVAEVFTGSPGKYVGLM    |
| 15por | STLQRYKELQDIIAILGLDELSEEDRLIVSRARKIERFLSQPFFVAEVFTGSPGKYVSL     |
| 16syn | STLQRYKELQDIIAILGLDELSEEDRLIVDRARKIERFLSQPFFVAEVFTGAPGKYVSLA    |
|       |                                                                 |
| 01can | DSIKGFNMILSGEVDNIPEQAFYLVGRIEEAIKAKQVEKS-----MALTVKVITPDR       |
| 02chl | ETIQGFNLILSGELDDLPEQAFYLVGNLDEAVSKAATLS-----MTLQVCIMTPDR        |
| 03cya | DTTKGFTMILGDELDELPEQAFYLVGDIQEAISKGQKLAEAK-----MSLDVRVIAPNK     |
| 04eug | ETIEGFKAISGLDEVPEQAFYLVGNIEEVISKAIELQ-----MTLDVSIIPER           |
| 05gui | DTMKGFNMILNGELDELPEQAFYLVGNIDEAIEKAKSLK-----MSIHISIIAPDR        |
| 06mai | ETIRGFQLILSGELDGLPEQAFYLVGNIDEASTKAINLEESKLK-----MKLNLVLTTPKR   |
| 07mar | ETIKGFQMILSGELDSLPEQAFYLVGNIDEATAKAATLQVES-----MLNLRIAPNR       |
| 08mes | DSIKGFTMILNGELDSLPEQAFYLVGNIEEAIKAAATLKE-----MSFQVRIIAPNG       |
| 09nep | ESIQGFKLILSGELDSLPEAFYLVGNIEEAIKAAASIAAAAAAAMSLQVRILTPEQ        |
| 10nic | ETIRGFQLILSGELDGLPEQAFYLVGNIDEATAKAMNLEMESLKK-----MTLNLVLTTPNR  |
| 11odo | ETIKGFTMILNGELDELPEQAFYLVGNIDEAIAKAETLK-----MVMNVRVLTPT         |
| 12oen | ETIRGFQLILSGELDGLPEQAFYLVGTIDEATAKANLEMESLKK-----MTLNLVLTTPNR   |
| 13ory | ETIRGFQLILSGELDGLPEQAFYLVGNIDEASTKAINLEENKLK-----MKLNLVLTTPKR   |
| 14pin | ETIRGFQMILSGELDGLIEQSFYLVGNIDEATAKAINSSMES-----MTLNLVLTSPNR     |
| 15por | DATKGFQMLKGLDELDELPEQAFYLVGDIQEAISKADSKMD-----MTLNIRIIPDR       |
| 16syn | DTTKGFKAISLDELDDLPEQAFYLVGDIIEEAKAKGAKLKEG-----MTLTVRVTTPDK     |
|       |                                                                 |
| 01can | VVWKKTVEEIILPSSTGQLGILMNHAPLLTALDIGVMRARM---VNTWVPLVLLGGFAQI    |
| 02chl | IFWNDQADEIILPTNTGQMGVLTNHPALITALDIGVTILR---S-NSNWNVALMGGFALV    |
| 03cya | VIWAKNAEEVILPSQSGMLGILTSHAPLYTALNTGVMKIRN---ETGWTISVVMGGFVEV    |
| 04eug | VFWKRVEEIIILPTLSGQMGVLKDHIPILTGLDIGILVRQKS--SSDWTSLVVTGGFALI    |
| 05gui | IIVWDANAEEVILPSSTGQLGILGHAPLLTALDIGVMRVRV---DRDWTPIVLLGGFAEI    |
| 06mai | IIVWCEVKEIILSTNSGGIGVLPNHAPINTAVDMGPLRIRL---LNDQWLTAVLWSGFARI   |
| 07mar | IIVWSDIQEIIILSTNSGGIGILPNHASVLTALDIGIVKIRL---NDQWSTMALMGGFAMI   |
| 08mes | IIVWSEAEIILSTNTGKIGVLTNHTSLTGLDIGTIRIRA---LNNQWNTALMSGFAMI      |
| 09nep | VFLNVSADIIILPTNTGQMGVLTNHTPLITADIGPMLVR---S-ESTWQSMALLGGLALV    |
| 10nic | IIVWSEVEEIVLSTNSGGIGILPNHAPIATAVDIGILRIRL---NDQWLTALMGGFARI     |
| 11odo | VICSTTADDEVILPGLTGLVGILDGHAALITALDTGLLRIRL---NEKWTPILCGGLAEI    |
| 12oen | IIVWSEVKEIILSTNSGGIGVLPNHAPIATAVDIGILRIRL---NGQWLTALMGGFARI     |
| 13ory | IIVWCEVKEIILSTNSGGIGVLPNHAPINTAVDMGPLRIRL---LNDQWLTAVLWSGFARI   |
| 14pin | IIVWSEVKEIILSTNSGGIGVLPNHASVAAVDIGVMKIRL---NGQWSTMMMGFAKI       |
| 15por | IIVWDAAEQEIIILPSSTGQLGILTGHAFLTALDIGVMRVRV---DKEWMPILVLLGGFAEI  |
| 16syn | VVWDEEVQELILPSTTGQLGILSNHAPLLTALDIGVMRVRP---GKDWNIAVMGGFAEV     |
|       |                                                                 |
| 01can | DNNLVTTIIVSDAEVKAIDEEANKLLAASLANMEKAQSNREK---IESMQNLRRARARM     |
| 02chl | KQNQVTILVNEAESAQTIQVDEAEIAFQEAQKLEQSQG-EKQR---VEATFVFKRARARY    |
| 03cya | EKNEVLVLVNAAGEYVDEIDLAAKDKVEKALETFNSAEAPKEK---EEAAEFKYAQAARL    |
| 04eug | NSNNVTILVNEAEFGSEINVEQAQISYNSSSKHALEMNKD-IKRR---FELTNLKKARARF   |
| 05gui | ENDELTILVNGAEAEASQIDRDQAQRDELEMTVKFNEATTNKR---IEATQNLRKARARL    |
| 06mai | VNNEIILGNDDELGSIDPEEAQQALETAEANLSKAEG-TKEL---VEAKLALRRARIRV     |
| 07mar | DNNNLITILVNDAEKASEIDYQEAQETFPKAKNTLEEAEGNKKE---IEALLVFKRAKARL   |
| 08mes | KDNVATIIIVSEANGANIKSEEAQTLQEEAKSYFNTAKG-TKNE---VEANLAVKRAETRL   |
| 09nep | KDNQVILVNEAELGSDINAEAEATTFLAAKEALANSKT-RKQ---IENNLAFKRARVRY     |
| 10nic | GNEITVLVNDAEKGSIDPEQAQQTLELAEANVKAEG-RRQK---IEANLALRRARTRV      |
| 11odo | DRNRVTVLVNDVEELVAVELNEATTELEKATLAVENAETSKAR---LDASIELKKAVARL    |
| 12oen | GNEITVLVNDAEKGSIDPEQAQETLGLAEANFRKAEG-KRQT---IEANLALRRARTRV     |
| 13ory | VNNEIILGNDDELGSIDPEEAQQALETAEANVSRAEG-TKEL---VEAKVALRRARIRV     |
| 14pin | DSDRITILVNNAAERDVIDPREAQENFRIAKADLARAEG-KRQA---IEADLALKARARL    |
| 15por | ENNQVILVNGAEAEASQIDLVEAKNLDATQLLNDASSSKEK---IEATQNIKKARARV      |
| 16syn | ENNEVKVLVNGAELGTTIDAEASARQYTTAAQGALEANRGDKPNKLASNNYKARARL       |

01can QAILSLK-----MRSQSLNFVASNTT---SDFISNLLSESVINIIILISLLIY  
02chl QVVKQLGV-----MLLSTSFLLAGH---FGFNTNLETNVNLAVVLAIVLT  
03cya KAVVDK-----MNLINSTKIFSSLMIASVENNTAIFSLNTDILETNILNLLVIFFLIY  
04eug QVTQLKK-----MVIDNFNIIFTIISNAKT---FGINTNVFETNINLAIVVGTLFY  
05gui QAVSA-----MDIISGFYNTINLAELSNAKT---FGFNPNILEANVNLIAILLSGVIY  
06mai EAVNWIPPSN----MKNVTHSFVFLAHWPFAGS---FGLNTDILATNINLTVVVGVLIF  
07mar EAINMASKL----MENGTYFIISSNFWTIAGS---FGLNTNLETNINLGVVGLLVY  
08mes KASQNL-----MTDILTNIFTILSRMSLAEG---FGFNTDIFETNINLAVVGLILT  
09nep QVATLVK-----MFHFLALTPLAHSEG---FGLNTNILETNINLAAVFALLAY  
10nic EAINPIS-----MKNVTSFVSLGHWPASGS---FGFNTDILATNPINLSVVLGVLIY  
11odo EGMNYS-----MENFNQIFT---LLAENEG---IGLNTDILETGIINIAALVGILY  
12oen EAINVIS-----MKNVTSFVSLVHWPSAGS---FGFNTDILATNPINLSVVLGVLIY  
13ory EAVNWIPPSN----MKNVTHSFVFLAHWPASGS---FGLNTDILATNINLTVVVGVLII  
14pin EAINASPPVSN---MKNVIDPFISLYWPSAGG---FGSNTNILETNINSSVVLVLIY  
15por QAATAASA---MNSIVNIPTIIILSEHSSEHT---FGFNSDIFEANVINILLFLGLIY  
16syn QAAGGAV-----MLNTFLIAAEAEHEAGEGG---FGINLDFLEANFNLAILLGIITY

01can LGKRTIVTNLNRKLNQCSILEAEKLSQAQLRLAEAEKKITESEIIILKHEAKIAA  
02chl YVGDALRGLLANRKQSILTNFREADQRATEAQNKLREAEQLEQAQAKQIREQANVTI  
03cya QGRPFPTALLEERRKTVLDKIKKSENSYNEAEKLEKAKSLEQAELAAQIYEEAEAVA  
04eug YGKLTSLDLKTRKKTIIKNILDIDEKIRSSQSLLYLAELFEFNAAKKASLRISNGTTF  
05gui LGRNFLTSALESRQKVTAEIQEAERLQANVKLLDAEQQLTQQTQVIEQIKKEAEKTA  
06mai FGKGVLDLLNRKQRIILSTIRNSELKRGTLQLEKARIRLQKVELEADEYRMNGYSEI  
07mar FGKGVLSNLLNRRKLTILNTIQDAEERYKEATDKLNQARTLQQAQKADDIRINGLSQM  
08mes SGREFVSVLLQNRQQNIIQSINDADERYKEAEKLLQAQNEFEQAKLEADQILAQSKKTA  
09nep VGTDFVSSLLKTRKESILKSLRDADERYQDAVNQLKQALQELTARTNAEIRRGSEINA  
10nic FGKGVLDLLNRKQRIILSTIRNSELKRGTAIEQLEKARSRLRKVSEAEQFRVNGYSEI  
11odo AGRDFLGSFLEQRKTSIVQGVQDAEGRLEENRRLSEAQQLSQAHVISEIKNETISAK  
12oen FGKGVLSNLLNRRKQRIILSTIRNSELREGAIEQLEKARARLDQVQIEAEGYRAYGYFGI  
13ory FGKGVLDLLNRKQRIILSTIRNSELRRGTIEQLEKARIRLQKVELEADEYRMNGYSEI  
14pin FGKGVLSNLLNRRKQRIILSTIRNSELCKGAIDQLEKARACLRNVMIADEIQVNGNSQI  
15por VIKQSLGSTLNRQLKVLAAIQESEERLEQASSRLESEKQLAQGTQIIINQKKEAQLTA  
16syn YAPKTLGKILGDRRQKIADAIEEAETRQRKSAQILAEKKLAQAKAEAAIRVQEAQGRA

01can NKIKELIVNNAKRKIDKILDGKTTIANTEIEMNQIKNRIISTAIDKTKIKLTKNLQ-A  
02chl EQEKKQFIRQTQEDIKRLGLTQQETLKFQEQKAQNELAEKLVKALQVQREKLNRLT-S  
03cya ESIKKTGLAQLEKDIKRIEETQASINTQQLSVITYLRQVALLALRRVVSQKLNLYK-P  
04eug LKSFDIIRSSVNEDIKRLKQSKRLILRTEDKKSVEIFKNLYSQAQKAKATIIKRLN-S  
05gui RTVKETILAQGKLDIERLTNNGKSSIEKAEQIKKKIQQHTDLAIKKVSAQMETFMT-D  
06mai EREKENLINATSIQLEKSKNETLYFEKQAMNQVRQQGFQQAQVQALGTLSCLN-T  
07mar EKEKQDLINAADEDKSKRLSDSKNATIRFEKQRAIEQVRQVQVSRALERALETLKSRLN-S  
08mes SEIEVGLMNLIKEDTKLLDMKQATISFEEKAISEIRRQVIRLALQRALEQSKSRLN-R  
09nep EAIRQRLELLTQEEMARLEEAKETIIKLEEEKAAEVCTKVISMALVRAEKTISSMD-E  
10nic EREKLNINSTYKTLQLENYKNETIQFEQQRINQVRQVQFQALRGALGTLSCLN-N  
11odo KVLLESDAYQAKKDLTTRFSRALATFRSKERQIFLEVKEIIILLVLKRTVARAQQTGPK  
12oen DEQRHESINSTYKTLQLENNKNSIHFEQQRINQVRQFQFQALQALGTLSCLN-N  
13ory EREKANLINATSIQLEKSKNETLYFEKQAMNQVRQVQFQQAQVQALGTLSCLN-T  
14pin EREKEDLLNTASDNLEQLEDPKNETIYSEQQRADFQIRQVSRQALRRAIGTLNRLN-T  
15por EYKRVSSILAQGGIDIERLAIITGKSNIEAEKQIRRRQIQQAIFALFKVTLQLENQMS-S  
16syn EVAKQETATQTEADLRRIEAAAQDLGAEQERVIAELKRRIAEQAVAKAEADLRDLN-E

01can ETIERITDNKITFLNNL----MDPIISAASVIAAGLAVGLAAIGPGIGQGSAAANAVE  
02chl SIHSAVNNFQIVLFTNSKAS---MNPIVAAASVIAAGLAVGLAAIGPMGQGTAAGYAVE  
03cya ELHSQFIDRKILQLRKQK----MDATVSAASVIAAALAVGLAAIGPGIGQGTAAAGQAVE  
04eug KIHKKIILKKMEKMSLKKLPKYMNPICAAVIGAGLAIGLGAIIGPGIGQGTASGKAIE  
05gui NLQVQVDTNIIASLGKGI----MNPIVSAASVIAAGLAVGLAAIGPGIGQGTAAAGQAVE  
06mai ELHFRTIRANIGILGAIEWKR--MNPLIAAASVIAAGLAVGLASIGPGVGGTAAAGQAVE  
07mar ELHLRIMDYHIGLLRAMESTIE--MNPLIAAASVIAAGLAVGLASIGPGIGQGTAAAGQAVE  
08mes RLQKRVTRLNIGLLRLVTPNDVMSPLISAASVIAAGLAVGLASIGPGVGGTAAAGQAVE  
09nep AMHRRVMDMYLMLLREVY----MSPLIAAASVIAAGLAVGLASIGPGIGQGTAAAGQAVE  
10nic ELHLRTIRSNIGMLGTMKETTD--MNPLISAASVIAAGLAVGLASIGPGVGGTAAAGQAVE  
11odo ERATALITETINKLEGDLL----MDSIIAASVIAAGLAIGLAAIGPGIGQGNAAAGQAVE  
12oen ELHLRTISANIGLFGSMKELTD--MNPLISAASVIAAGLAVGLASIGPGIGQGTAAAGQAVE  
13ory ELHFRTIRANISILGAME----MNPLIAAASVIAAGLAVGLASIGPGVGGTAAAGQAVE  
14pin ELHLRTIDHNIIGLLRTMMTND--MDPLISAASVIAAGLAVGLASIGPGVGGTAAAGQAVE  
15por DIQLRIIDNNIAKLGQDL----MDSIVSAASVIAAGLAVGLAAIGPGIGQGSAAANAVE  
16syn DTQDRLIERSIAQLGGR-----MDSTVAAASVIAAALAVGLAAIGPGIGQGNASGQAVS

01can GLARQPEAEGKIRGTLTLLSLAFMESLTIYGLVVALSLLFANPFK-----MSKFYDWF  
02chl GIARQPEAEGKIRGALLSLAFMESLTIYGLVVALALLFANPFAG-----MGKVYDWF  
03cya GIARQPEVDGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MSKVYDWF  
04eug GLARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALAIIFANPFV-----MSRVYDWF  
05gui GIARQPEAEGKIRGTLTLLSLAFMESLTIYGLVVALALLFANPFTS-----MGKVYDWF  
06mai GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MSKVYDWF  
07mar GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MGKVYDWF  
08mes GIARQPEAEGKIRGTLTLLSLAFMESLTIYGLVVALALLFANPFVS-----MSKVYDWF  
09nep GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFVS-----MSKIYDWF  
10nic GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MSKVYDWF  
11odo GIARQPEGENKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFG-----MGKVYDWF  
12oen GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MSKVYDWF  
13ory GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MSKVYDWF  
14pin GIARQPEAEGKIRGTLTLLSLAFMEALTIYGLVVALALLFANPFV-----MGKVYDWF  
15por GIARQPEVEGKIRGTLTLLSLAFMESLTIYGLVVALSLLFANPYVG-----MSKIYDWF  
16syn GIARQPEAEGKIRGTLTLLSLAFMESLTIYGLVVALALLFANPFA-MFSKEVTESKVFQWF

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01can | QERLEIQIADDISAKYVPPHNVFYCFGGMTLTCFLVQLATGFAMTFYKPTTIEAFSS       |
| 02chl | EERLEIQSIADDISKYVPPHNVIFYCLGGITFTCFLVQVATGFAMTFYRPTVAEAFAS      |
| 03cya | QERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFAS      |
| 04eug | EERLEIQAIADDISKYVPPHNVIFYCLGGITFTCFIIQVATGFAMTFYRPTVTEAFLS      |
| 05gui | EERLEIQAIADDISKYVPPHNVIFYCLGGITFTCFIIQVATGFAMTFYRPTVAEAFAS      |
| 06mai | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFSS      |
| 07mar | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFSS      |
| 08mes | NRLEIQGIADDISKYVPPHNVIFYCLGGITFTCFIMQVASFAMTFYRPTVTEAFAS        |
| 09nep | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFAS      |
| 10nic | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFAS      |
| 11odo | EERLEIQAIADDISKYVPPHNVIFYCFGGIVFTCFLVQVATGFAMTFYRPSVDAFAS       |
| 12oen | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFAS      |
| 13ory | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFSS      |
| 14pin | EERLEIQAIADDISKYVPPHNVIFYCLGGITLTCFLVQVATGFAMTFYRPTVTEAFAS      |
| 15por | EERLEIQAIADDISKYVPPHNVIFYCLGGIVFVSFLIQVATGFAMTFYRPTVAEAFAS      |
| 16syn | NRLEVQAISSDIASKYVPPHNVIFYCLGGITLTCFLIQFATGFAMTFYKPTVTEAFAS      |
|       |                                                                 |
| 01can | IQHIMTQVSFGWLIRSLHRWSASMMVLMILHIFRVYLTGGFKKPRELWTGVLAVLT        |
| 02chl | VQYIMTEVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVMIMAVLT       |
| 03cya | VQYIMTDVNFGLIRSTRHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVLAVIT        |
| 04eug | VKYIMNEVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGIIILAILT       |
| 05gui | VEYIMTEVNFGLFRSMHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILAVVT       |
| 06mai | VQYIMTEANFGWLIRSVHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTVLAVLT      |
| 07mar | VQYIMTEVNFGLIRSVHRWSASMMVLMILHIFRVYLTGGFKKPRELWTGVLAVLT         |
| 08mes | VQYIMTDVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILAVCT       |
| 09nep | VEYIMTNVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILAVIT       |
| 10nic | VQYIMTEANFGWLIRSVHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTVLAVLT      |
| 11odo | VEYIMTSVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILAVVT       |
| 12oen | VPIYIKTEANFGWLIRSVHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTVLAVLT     |
| 13ory | VQYIMTEANFGWLIRSVHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTVLAVLT      |
| 14pin | VQYIMTEVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILAVLT       |
| 15por | VEYIMTDVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTILGLVT       |
| 16syn | VQYIMNEVNFGLIRSIHRWSASMMVLMILHVCRVYLTGGFKKPRELWTGVTMLAVTT       |
|       |                                                                 |
| 01can | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVVGDNVVEILRGGTGVQATLTRFYSLHT     |
| 02chl | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVVGPAIVELLRGGSVGVQSTLTRFYSLHT    |
| 03cya | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSNVVELLRGSVSVGQSTLTRFYSLHT    |
| 04eug | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGNIVELLRGGSVSVGQSTLTRFYSLHT    |
| 05gui | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVVGSLVELLRGSVSVGQATLTRFYSHT      |
| 06mai | ASFSGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSASVGVQSTLTRFYSHT   |
| 07mar | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSVSVGQSTLTRFYSLHT    |
| 08mes | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGAPLVELLRGSVGVQSTLTRFYSLHT     |
| 09nep | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGAPLVELLRGSVSVGQSTLTRFYSHT     |
| 10nic | ASFSGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSASVGVQSTLTRFYSHT   |
| 11odo | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVVGPPVLVLRGGESVGVQSTLTRFYSHT     |
| 12oen | ASFSGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSASVGVQSTLTRFYSHT   |
| 13ory | ASFSGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSASVGVQSTLTRFYSHT   |
| 14pin | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVIGSPLVELLRGSVSVGQSTLTRFYSHT     |
| 15por | VSFGVTGYSLPWDQVGYWACKIVTGVPEAIPVVGESIVELLRGGSVGVQSTLTRFYSHT     |
| 16syn | VTFGVTGYSLPWDQVGYWACKIVSGVPEAIPVVGQDLVTLRGSESVGVQATLTRFYSLHT    |
|       |                                                                 |
| 01can | LFLPALSVIFLLAHFLMIRKQGISGPLMVEVLLTGIVLGSIFITLLGLLAAKLQYNR--     |
| 02chl | FVLPPLATAVFLMHFLMIRKQGISGPLMVEALLSGIVLGLPVTIAGLFVTAYLQYRR-G     |
| 03cya | FVLPPLLSAVFMLVHFLMIRKQGISGPLMVEPLLSGIVLGLIPVTILGLFVAAYLQYRR-G   |
| 04eug | FVLPPLLTATFMLGHFLMIRKQGISGPLMVEVLLSGIILGLIPITIGLFFTAYLQYMRSG    |
| 05gui | FVLPPLAAVFLMLTHFLMIRKQGISGPLMIEVLLFGIVLGLIPVTILGLFVAAYLQYRR-G   |
| 06mai | FVLPPLLTAVFMLMHFPMIRKQGISGPLMIEVLLFGIVLGLIPITILAGLFVTAYLQYRR-G  |
| 07mar | FVLPPLLTAVFMLMHFLMIRKQGISGPLMVEALLSGIVLGLIPITILGLFVTAYLQYRR-G   |
| 08mes | FVLPPLLTAVFMLAHFLMIRKQGISGPLMVEALLSGIVLGMIPITILAGLFVTAYLQYRR-G  |
| 09nep | FVLPPLLTAVFMLMHFLMIRKQGISGPLMVEALLSGIVLGLIPVTILAGLFVTAYLQYRR-G  |
| 10nic | FVLPPLLTAVFMLMHFPMIRKQGISGPLMIEVLLFGIVLGLIPITILAGLFVTAYLQYRR-G  |
| 11odo | FVLPPLAAVFLMLTHFLMIRKQGISGPLMVEPLLSGIVLGMITVSALGLFVAAYLQYRR-G   |
| 12oen | FVLPPLLTAVFMLMHFLMIRKQGISGPLMIEVLLFGIVLGLIPITILAGLFVTAYLQYRR-G  |
| 13ory | FVLPPLLTAVFMLMHFLMIRKQGISGPLMIEVLLFGIVLGLIPITILAGLFVTAYLQYRR-G  |
| 14pin | FVLPPLLTAVFMPMHFLMIRKQGISGPLMIEALPSGIVLGLIPITILAGLFVTAYSQYRR-G  |
| 15por | FVLPPLLTAVFMLMHFLMIRKQGISGPLMVEPLLSGIILGLIPVTILSGLLVAAYLQYRR-G  |
| 16syn | FVLPPLAAVFLMLTHFLMIRKQGISGPLMIEVLLSGIVLGLIPVTILAGLFVTAAYLQYRR-G |
|       |                                                                 |
| 01can | KKTN--MNQILKEKEQKTAKILVDRNPVPTSFEKWGPGHFSRSLAKG-PKTTTWIWNLH     |
| 02chl | DQLNI-MTISPPEREAKKIVVDRNPVATNFEKWAKPGHFSRSLAKG-PKTTTWIWNLH      |
| 03cya | NQFEF-MRISPPEREAKKIVIDKDPVSTSFEDKAVPGHFSRSLAKG-PKTTTWIWNLH      |
| 04eug | NSFY--MTITPPEQKVRVAFVSNPVTSFEKWSRPGHFSRSLAKG-PNTTTWIWNLH        |
| 05gui | NQFGL-MTISSTEQEAKKVVVDRNPVSTSFEDKAVPGHFSRSLAKG-PKTTTWIWNLH      |
| 06mai | DQLDL-MIRSEP---EVKIAVDRDPKTSFEKWARPGHFSRTIAKGNPDPTTWIWNLH       |
| 07mar | DQLDL-MTIRSEP---EVKIVVEKDPKTSFEKWARPGHFSRSLAKG-PSTTTWIWNLH      |
| 08mes | DQLKI-MTISPPEREAN-GKIVVDRDPKTSFEKWARPGHFSRSLAKG-PNTTTWIWNLH     |
| 09nep | DQLNL-MTISPPEREAKKIVVDRNPVSTSFEDKAVPGHFSRSLAKG-PTTTTIWDLH       |
| 10nic | DQLDL-MIRSEP---EVKILVDRDPKTSFEKWARPGHFSRTIAKGNPDPTTWIWNLH       |
| 11odo | NQFEI-MAISSTERRAKNVQIFVEKDAVETSAKWAQPGHFSRSLAKG-PKTTTWIWNLH     |
| 12oen | DQLDL-MIRSEP---EVKILVDRDPKTSFEKWARPGHFSRTIAKGNPDPTTWIWNLH       |
| 13ory | DQLDL-MIRSEP---EVKIVVDRDPKTSFEKWARPGHFSRSLAKG-PDPTTWIWNLH       |
| 14pin | DQLDI-MTIRSEPVEKVKVVDVDRDTPKTSFEKWARPGHFSRSLAKG-PDPTTWIWNLH     |
| 15por | NQLGL-MAISSKEQETKKVIVSDKNPVDTSFEKWARPGHFSRSLAKG-PKTTTWIWNLH     |
| 16syn | NQFNLDMTISPPEREAK-AKVSVDNPNVPTSFEKWARPGHFDRLARG-PKTTTWIWNLH     |

|       |                                                               |
|-------|---------------------------------------------------------------|
| 01can | ADVHDFDSHTSSLEDISRKFSAHFGQLSLIFLWLSGMYFHGARFSNYTIWLSNPSLIKP   |
| 02chl | ADAHDFDTQTSDEEISRKFVSAHFGQLGIIFLWLSGMYFHGARFSNYEAWLTDPTHIKP   |
| 03cya | ADVHDFDSYTSDEEDVSRKFSAHFGQLAVFVLWLSGMYFHGARFSNYEAWLSNPPTTIKP  |
| 04eug | ADAHDFDNHTTDLIEDISRKFVSAHFGQLAIIEIWLSCMFHGHARFSNYEAWLLDPIHVKP |
| 05gui | ADAHDFDSHTSSLEDISRKFSAHFGQLSIIFLWLSGMYFHGARFSNYSAWLSNPPTTVKP  |
| 06mai | ADAHDFDSHTGDLEEISRKFVSAHFGQLSIIFLWLSGMYFHGARFSNYEAWLSDPTHIGP  |
| 07mar | ADAHDFDSHTNDLEEISRKFVSAHFGQLAIIFLWLSGMYFHGARFSNYEAWLSDPTHIKP  |
| 08mes | ADAHDFDSHTNDLEDISRKFVSAHFGQLAVIFLWLSGMYFHGARFSNYEAWLSDPTHIKP  |
| 09nep | ADAHDFDSHTTDLIEDISRKFSAHFGQLGVILWLSGMYFHGARFSNYEAWLSDPTHIKP   |
| 10nic | ADAHDFDSHTSDLEEISRKFVSAHFGQLSIIFLWLSGMYFHGARFSNYEAWLSDPTHIGP  |
| 11odo | ADAHDFDSQTSLSLEEVSRKFSAHFGQLAVIFLWLSGMYFHGAIFYSNYSAWLSDPISIKQ |
| 12oen | ADAHDFDSHTSDVEAISRKVSSAHFGQLSLIFLWLSGMYWGFALFSNYEAWLSDPTHIGP  |
| 13ory | ADAHDFDSHTGDLEEISRKFVSAHFGQLSIIFLWLSGMYFHGARFSNYEAWLSDPTHIGP  |
| 14pin | ADAHDFDSHTNLEDISRKFSAHFGQLAIIFLWLSGMYFHGARFSNYEAWLADPTHIKP    |
| 15por | ADAHDFDSQTSLSLEEVSRKFSAHFGQLAVIFLWLSGMYFHGARFSNYVAWLSNPPTGIKP |
| 16syn | ANAHDFDSQTSLEDVSRKFSAHFGQLAVFVWLWLSGMYFHGAKFSNYEGWLADPTHIKP   |
|       |                                                               |
| 01can | SAQIVWPIVGQEILNADLGGGSGGIQITSGFFHLWRASGITNELELYVTALGGLFMAGLM  |
| 02chl | SAQVWVPIVGQEILNADVGGGFQGLQITSGFFQLWRASGITSELQLYTTAIGGLVMAAAM  |
| 03cya | SAQVWVPIVGQEILNGDVGGGFQGIQITSGFLQMWASGITTELQLYVTAIGALVMAALM   |
| 04eug | SAQIVWPIVGQEILNGDVGGGFQGIQITSGFLQLWRSGGITSEFQLYTTALTGLIFSAYL  |
| 05gui | SAQVWVPIVGQEILNGDVGGGFQGVQVITSGFFQIWRASGITSEVELYWCVAAGLLMSGLM |
| 06mai | SAQVWVPIVGQEILNGDVGGGFRGIQITSGFFQIWRASGITSELQLYCTAIGALIFASLM  |
| 07mar | SAQVWVPIVGQEILNGDVGGGFQGIQITSGFFQLWRASGITSELQLYSTAIGGLVFAALM  |
| 08mes | SAQVWVPIVGQEILNGDVGGGFQGVQITSGFFQLWRASGITVNEQQLYTTAIGGLIAAGLM |
| 09nep | SAQVWVPIVGQEILNGDVGGGFQGIQITSGFFQLWRASGITSELQLYSTAIGGLLAAAM   |
| 10nic | SAQVWVPIVGQEILNGDVGGGFRGIQITSGFFQIWRASGITSELQLYCTAIGALVFAALM  |
| 11odo | SSQVWVPIVGQEILNADVGGNFQGVQITSGWFQMWASGITSEVELYWIAIGGLIMSALM   |
| 12oen | SAQVWVPIVGQEILNGDVGGGFRGIQITSGFFQLWRASGITSELQLYCTAIGALIFAALM  |
| 13ory | SAQVWVPIVGQEILNGDVGGGFRGIQITSGFFQIWRASGITSELQLYCTAIGALIFASLM  |
| 14pin | SAQIVWPIVGQEILNGDVGGGFRGIQITSGFFQIWRASGITSELQLYCTAIGALIFAALM  |
| 15por | SAQVWVPIVGQEILNGDVGGGFQGVQITSGWFQLWRASGITTEFQLYCTAIGGLMAALM   |
| 16syn | SAQVWVPIVGQILNGDVGGGFHGIQITSGFLYLWRASGFTDSYQLYCTAIGGLVMAALM   |
|       |                                                               |
| 01can | AFCGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINKLLDAGVD    |
| 02chl | FFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINKLLDAGVD    |
| 03cya | LFAGWPHYHKAAPKLEWFQNAESMMNHHLAGLLGLSLWAGHQIHVSLPVNKLDSGVS     |
| 04eug | FFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINKLLDSGVN    |
| 05gui | IFAGWPHYHKAAPKLEWFQNAESMLNHHLAGLLGLSLWAGHQIHVSLPINKLLDAGVA    |
| 06mai | LFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLDAGVD     |
| 07mar | LFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLDAGVD     |
| 08mes | FFAGWPHYHKAAPKLEWFQNAESMMNHHLAGLLGLSLWAGHQIHVSLPVNQLDAGVD     |
| 09nep | FFAGWPHYHKAAPKLEWFQNVESMMNHHLAGLLGLSLWAGHQIHVSLPVNKLNAGVD     |
| 10nic | LFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLNAGVD     |
| 11odo | LFAGWPHYHKAAPKLEWFQNAESMMNHHLAGLLGLSLWAGHQIHIALPINQLDAGVA     |
| 12oen | LFAGWPHYHKAAPKLEWFQNVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLNAGVD     |
| 13ory | LFAGWPHYHKAAPKLEWSDVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLDAGVD      |
| 14pin | LFAGWPHYHKAAPKLEWFQEVESMLNHHLAGLLGLSLWAGHQIHVSLPINQLDAGVD     |
| 15por | LFAGWPHYHKAAPKLEWFQNVESMMNHHLAGLLGLSLWAGHQIHVSLPINKLDSGVS     |
| 16syn | LFAGWPHYHKAAPKLEWFQNVESMMNHHLAGLLGLSLWAGHQIHVSMPINKLLDAGVA    |
|       |                                                               |
| 01can | PKTIPLPHEFILNRELMSQLYPSFEKGLWPFFSLNWGAYSDFLTFRGGLNPVTGSLWMSD  |
| 02chl | PKEIPLPHEFLNPPELMAQLYPSFAKGLAPFFTLNWAQYSDFLTFQGGGLNPVTGGLWLT  |
| 03cya | PQEIPLPHEFILNKDLIAQLYPSFGQGLTPFFTLNWNEYSDFLTFKGGGLNPVTGGLWLS  |
| 04eug | PAELPLPHDFILDKSLISQLYPSFSKGLAPFFTFHWSEYSDFLTFRGGGLNPVTGGLWLT  |
| 05gui | PQEIPLPHEFILNMRDLMAQLYPSFNKGLIPFFSLNWSEYSDFLTFKGGGLNPVTGGLWLS |
| 06mai | PKEIPLPHEFILNRDLAQLYPSFAEGATPFFTLNWSKYAEFLSFRGGLDPIITGGLWLS   |
| 07mar | PKEIPLPHEFILNRDLAELYPSFAKGLTPFFTLNWSEYSDFLTFRGGGLNPVTGGLWLT   |
| 08mes | PKEIPLPHEFVMNRELMAQLYPSFAKGLAPFFTLNWGEYSDFLTFRGGGLNPVTGGLWLT  |
| 09nep | PKEIPLPHEFLNRDLMAQLYPSFAKGLTPFFTLNWAEGDFLTFRGGGLNPVTGGLWLS    |
| 10nic | PKEIPLPHEFILNRDLAQLYPSFAEGATPFFTLNWSKYADFLTFRGGGLDPVTGGLWLT   |
| 11odo | PQEIPLPHEFILNRELMAQLYPSFNKGLAPFFSGQWGEYSDFLTFKGGGLNPVTGGLWLS  |
| 12oen | PKEIPLPHEFILNRDLAQLYPSFAEGATPFFTLNWSKYAEFLTFRGGGLDPVTGGLWLT   |
| 13ory | PKEIPLPHEFILNRDLAQLYPSFAERATPFFTLNWSKYAEFLSFRGGLDPIITGGLWLS   |
| 14pin | PKEIPLPHEFILNRDLAQLYPSFAKGLTPFFTLNWSKYAEFLTFRGGGLNPVTGGLWLT   |
| 15por | PQEIPLPHEFILNRELMAQLYPSFSKGLVPFFTLNWAESYDFLTFKGGGLNPVTGGLWLS  |
| 16syn | PKDIPLPHEFILPSKMAELYPSFAQGLTPFFTLNWGVYSDFLTFKGGGLNPVTGGLWLS   |
|       |                                                               |
| 01can | IAHHHLAISVLFIVAGHMYRTNWGIGHSEIKELDAHRGPLTGSCHRGLEYALTTSWHANL  |
| 02chl | TVHHHLAIAVLFIVAGHQYRTNWGIGSSLKEILEAHKGPFTGEGHKGLYEILTTSWHAQL  |
| 03cya | RAHHHLAIAVLFIVAGHMYRTNWGIGHSMKEILETHKGPFTGEGHKGLYEIFTNSWHAQL  |
| 04eug | VAHHHLALAVLFILAGHMYRTNWGIGHDKIGLLESHTGPFTGQGHKGLYEIFTNSWHAQL  |
| 05gui | TAHHHLALAVLFIVAGHMYRTNWGIGHSMKEILEAHKGPFTGEGHKGLYEILTTSWHSQ   |
| 06mai | IAHHHLAIAILFLIAGHMYRTNWGIGHGLKDILEAHKGPFTGQGHKGLYEILTTSWHAQL  |
| 07mar | TAHHHLAIAVLFIVAGHMYRTNWGIGHSEIKELAHKGPFTGEGHKGLYEILTTSWHAQL   |
| 08mes | TVHHHVAIAVLFIVAGHMYRTNWGIGHSMKEILEAHKGPFTGEGHKGLYEILTTSWHAQL  |
| 09nep | TAHHHVAIAVLFIVAGHMYRTNWGIGHSMKEILEAHKGPFTGEGHKGLYEILTTSWHAQL  |
| 10nic | IAHHHLAIAILFLIAGHMYRTNWGIGHGLKDILEAHKGPFTGQGHKGLYEILTTSWHAQL  |
| 11odo | IAHHHLAISVLFIVAGHMYRTNWGIGHSMKEILEAHKGPFTGEGHKGLYEILTTSWHAQL  |
| 12oen | IAHHHLAIAILFLIAGHMYRTNWGIGHGLKDILEAHKGPFTGQGHKGLYEILTTSWHAQL  |
| 13ory | IAHHHLAIAILFLIAGHMYRTNWGIGHGLKDILEAHKGPFTGQGHKGLYEILTTSWHAQL  |
| 14pin | TAHHHLAIAVLFILAGHMYRTNWGIGHGLKDILEAHKGPFTGEGHKGLYEILTTSWHAQL  |
| 15por | TAHHHLALAVFLAAGHMYRTNWGIGHSMKEILEAHKGPFTGNGHEGLEYILTTSWHAQL   |
| 16syn | TAHHHLAIAVLFILAGHMYRTNWGIGHSMKEILEAHKGPFTGEGHKGLYEILTTSWHAQL  |



01can AINLALFGSLSIIVAHMHYAMPYPYISIDYPTQLSLFTHHTWIGGFCIVGASAHGAIFM  
02chl AINLALFGSLSIIVSHHMYAMPYPYLATDYGTQLSLFTHHMWIGGFCIVGAGAHAGIFM  
03cya SLNLALFGSLSIIVAHMHYSMPYPYLATDYATSLCLFTHHVWIGGFLIVGAGAAHAIIFM  
04eug SLNLAMMGSLSIIVAQHMYSMPPYPYIAIDYGTLSLFTHHVWIGGFCIVGAAHAAIIFM  
05gui AINLAMMGSLSIIIVAHMHYAMPYPFIATDYPTQLSIFTHHMWIGGFCIVGAAHAGIFM  
06mai SLNLAMLGSTTIIVAHMHYSMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
07mar ALNLAMLGSLTIIIVAHMHYAMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
08mes GLNLAMGSLSIIIVAHMHYAMPYPYLATDYGTQLSLFTHHMWIGGFCIVGAAHAAIIFM  
09nep AINLALFGSLSIIVAHMHYAMPYPYLATDYGTQLSLFTHHMWIGGFCVVGAAHAAIIFM  
10nic SLNLAMLGSLTIIIVAHMHYSMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
11odo AINLAMMGSLSIIIVAHMHYAMPYPYIATDYATQLSLFTHHMWIGGFCVVGAAHGAIFM  
12oen SLNLAMLGSLTIIIVAHMHYSMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
13ory SLNLAMLGSTTIIVAHMHYSMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
14pin AVNLAMLGSLTIIIVAHMHYSMPYPYLATDYGTQLSLFTHHMWIGGFLIVGAAHAAIIFM  
15por AINLAMMGSLSIIIVAHMHYAMPYPYIATDYPTQLSLFTHHMWIGGFCVVGAGAHASIFM  
16syn AINLALLGSLTIIIVAHMHYAMPYPYQIADYATQLSLFTHHMWIGGFLIVGAGAHGAIFM

01can IRDYVPSQHYNNVLDRLIRHRDALISHLNWVCIFLGTHSFGLYIHNDTMRALGRSQDMFS  
02chl VRDYDPTNNYNNLLDRVLRRHDAMIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
03cya VRDYDPAQNYNNLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMRALGRPQDMFS  
04eug VRDYDPALNFNNLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
05gui VRDYNPAQNYNNLLDRVIRHRDAIIISHLNWVCIFLGHSFGLYIHNDTMRALGRTQDMFS  
06mai VRDYDPTTRYNDLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
07mar VRDYDPTTYNNLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
08mes VRDYDPTNNYNNLLDRVIRHRDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
09nep VRDYDPTNNYNNLLDRVIRHRDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
10nic VRDYDPTTRYNDLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
11odo VRDYTPANNYNNLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMRALGRPQDMFS  
12oen VRDYDPTTRYNDLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
13ory VRDYDPTTRYNDLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
14pin VRDYDPTTYNNLLDRVLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMSALGRPQDMFS  
15por VRDYNPAENYNNLLDRIRHRDAIVSHLNWVCIFLGHSFGLYIHNDTMRALGRSQDMFS  
16syn VRDYDPAKNVNNLLDRMLRRHDAIIISHLNWVCIFLGHSFGLYIHNDTMRALGRPQDMFS

01can DTAIQLQPFAQWQISLHTLAPANTAPNALATTSYVFGG-EVVAVANKIAMPMKLGTD  
02chl DTAIQLQPVFAQWQVNTHTFLAPGFTAPNALASTSPSWG-DVAVGGKVAMMPISLGTAD  
03cya DAAIQLQPFAQWQVGNSSAAGNTAPNALRNASYAFGG-DIVSVGEKVAMMPISLGTAD  
04eug DSAIQLQPFAQWQVNTHTLAPTLTAFNLVSPPTPVWGG-DVVSISGKVAMMPIKLGTD  
05gui DTAIQLQPFAQWQVNIHTIAPGNTTPNALATASYAFGG-DVIVSGNKVAMMPISLGTAD  
06mai DAAIQLQPFAQWQVNIHAGAPVTAPGATTSTSLTWGGGELVAIGKVALLPIPLGTAD  
07mar DTAIQLQPFAQWQVNTHTALAPNTAPNALASTSLTWGGGDVAVGSKVALLPIPLGTAD  
08mes DTAIQLQPFAQWQVNRNLYAPGFSAPNALASSSAVWGG-DVAVGGKVAMMPIQLGTSD  
09nep DTAIQLQPFAQWQVNIHKTALAPGLTAPNALASTSPSWG-DVAVGGKVAMMPISLGTAD  
10nic DTAIQLQPFAQWQVNTHTALAPGATAPGATASTSLTWGGGDLVAVGKVALLPIPLGTAD  
11odo DKAIQLQPFAQWQVNIHLLAPGTTAPNALATTSYAFGG-DVVEVGGKIAMMPIKLGTD  
12oen DTAIQLQPFAQWQVNTHTALAPGATAPGATTSTSLAWGGGDLVAVGKVALLPIPLGTAD  
13ory DTAIQLQPFAQWQVNLHAGAPRLTAPGATTSTSLTWGGGELVAVGKVALLPIPLGTAD  
14pin DTAIQLQPFAQWQVNTHTASAPGSTAPGATASTSLTWGGGDLVTVGSKVALLPIPLGTAD  
15por DTAIQLQPFAQWQVSIHTLAPGNTAPNALATASYAFGG-EVVSVGNKVAMMPISLGTAD  
16syn DTAIQLQPFAQWQVLSLHTLAPGATAPNALATASYAFGG-ETIAVAGKVAMMPITLGTAD

01can FMVHHIHAFTIHVTLLILLKGVLFARNSRLIPDKANLGRFRPCDGPGRGGTCQVSSWDHV  
02chl FMVHHIHAFTIHVTLLILLKGVLYARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
03cya FLVHHIHAFTIHVTLLILLKGVLFARNSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
04eug FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
05gui FMVHHIHAFTIHVTLLILLKGVLFARNSRLIPDKANLGRFRPCDGPGRGGTCQSSAWDSV  
06mai FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
07mar FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
08mes FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
09nep FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
10nic FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
11odo FMVHHIHAFTIHVTLLILLKGVLYARSSRLIPDKANLGRFRPCDGPGRGGTCQSSWDHV  
12oen FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
13ory FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
14pin FLVHHIHAFTIHVTLLILLKGVLFARSSRLIPDKANLGRFRPCDGPGRGGTCQVSAWDHV  
15por FMVHHIHAFTIHVTLLILVKGLFARNSRLIPDKANLGRFRPCDGPGRGGTCQVSGWDHV  
16syn FMVHHIHAFTIHVTALILLKGVLYARSSRLIPDKANLGRFRPCDGPGRGGTCQVSGWDHV

01can FLGLFWMYNCISVVIHFHFSWKMQSDVWGTVSQNSLVSHVVGNGFSQSAITINGWLRDLFLW  
02chl FLGLFWMYNCISVVIHFHFSWKMQSDVWGTVSANG-VSHITGNGFAQSANTINGWLRDLFLW  
03cya FLGLFWMYNSLVVLFHFHFSWKMQSDVWGNVTADGAVSHITGNGFAQSSITINGWLRDLFLW  
04eug FLGLFWMYNCISVAIFHFHFSWKMQSDVWGTVLANK-VSHITGNGFSQGSITINGWLRDLFLW  
05gui FLGLFWMYNCISVVIHFHFSWKMQSDVWGTQVNDGTVTHTITGNGFAQSSITINGWLRDLFLW  
06mai FLGLFWMYNCISVVIHFHFSWKMQSDVWGTISDQGIVTHITGNGFAQSSITINGWLRDLFLW  
07mar FLGLFWMYNCISVVIHFHFSWKMQSDVWGTISEQGVVTHITGNGFAQSAITINGWLRDLFLW  
08mes FLGLFWMYNCLSVIHFHFSWKMQSDVWGSVTAQG-VSHITGNGFAQSANTINGWLRDLFLW  
09nep FLGLFWMYNCISVVIHFHFSWKMQSDVWGNVTAQG-VSHITGNGFALSNTINGWLRDLFLW  
10nic FLGLFWMYNAISVVIHFHFSWKMQSDVWGSVSDQGVVTHITGNGFAQSSITINGWLRDLFLW  
11odo FLGLFWMYNCISVVIHFHFSWKMQSDVWGTTIPDGNISHITGNGFAQSSITINGWLRDLFLW  
12oen FLGLFWMYNAISVVIHFHFSWKMQSDVWGISDQGVVTHITGNGFAQSSITINGWLRDLFLW  
13ory FLGLFWMYNCISVVIHFHFSWKMQSDVWGTISDQGVVTHITGNGFAQSSITINGWLRDLFLW  
14pin FLGLFWMYNAISVVIHFHFSWKMQSDVWGNISDQGVVTHITGNGFAQSSITINGWLRDLFLW  
15por FLGLFWMYNSLSVAIFHFHFSWKMQSDVWGSVSPSGNVSHITGNGFAQSAITINGWLRDLFLW  
16syn FLGLFWMYNSLSIVIHFHFSWKMQSDVWGTVSPDGSVTHVTLNFAQSAITINGWLRDLFLW

|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01can | AQASQVIQSYGSSISAYGLMFLAAHFIWAFSLMFLFS-GRGYWQELIESIVWAHNSKLKIA   |
| 02chl | AQSSQVIQSYGSALSAYGLIFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 03cya | AQASQVIQSYGSALSAYGLMFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKFA    |
| 04eug | AQSSQVIQSYGSPLSAYGLMFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 05gui | AEASQVIQSYGSALSAYGLIFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLHFA    |
| 06mai | AQASQVIQSYGSSLSAYGLFFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 07mar | AQASQVIQSYGSSLSAYGLLFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 08mes | AQASQVIQSYGSALSAYGLMFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 09nep | AQASQVIQSYGSALSAYGLIFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 10nic | AQASQVIQSYGSSLSAYGLFFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 11odo | SQASQVIQSYGSASSAYGLIFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLNFA    |
| 12oen | AQASQVIQSYGSSLSAYGLFFLGAHFVWAFSLMFLFSCGRGYWQELIESIVWAHNKLKVA    |
| 13ory | AQASQVIQSYGSSLSAYGLFFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 14pin | AQASQVIQSYGSSLSAYGLLFLGAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLKVA    |
| 15por | AQASQVIQSYGSALSAYGLIFLAAHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKIKVA    |
| 16syn | AQAANVINSYGSALSAYGLIMFLAGHFVWAFSLMFLFS-GRGYWQELIESIVWAHNKLNV    |
|       |                                                                 |
| 01can | PTIQPRALSITQGRAVGAAYHLLGGIATTWAFFLSRAISIGMATKFPKFSQSLSSDPTTR    |
| 02chl | PAIQPRALSITQGRAVGAAYHLLGGIATTWFFLARIILAVGMATKFPKFSQALAQDPTTR    |
| 03cya | PSIQPRALSITQGRAVGAAYHLLGGIATTWFFHARIISVGMATKFPKFSQALAQDPTTR     |
| 04eug | PNIQPRALSITQGRAVGAAYHLLGGIATTWFFLARIISVGMATKFPKFSQGLAQDPTTR     |
| 05gui | PAIQPRALSITQGRAVGLAHYLLGGIGTTWAFFLARIISVGMATKFPKFSQALAQDPATR    |
| 06mai | PATQPRALSIIQGRAVGVTHYLLGGIATTWAFFLARIIAVGMELRFPKFSQGLAQDPTTR    |
| 07mar | PAIQPRALSITQGRAVGAAYHLLGGIATTWAFFLARIIAVGMASRFPKFSQGLSQDPTTR    |
| 08mes | PAIQPRALSITQGRAVGAAYHLLGGIATTWAFFLARIIAVGMATKFPKFSQGLAQDPTTR    |
| 09nep | PSIQPRALSITQGRAVGAAYHLLGGIATTWAFFLARIIAVGMATKFPKFSQGLAEDPTTR    |
| 10nic | PATQPRALSIIQGRAVGVTHYLLGGIATTWAFFLARIIAVGMALRFPKFSQGLAQDPTTR    |
| 11odo | PAIQPRALSITQGRAVGLAHYLLGGIGTTWFFLARIISITMATKFPKFSQALAQDPATR     |
| 12oen | PATQPRALSIVQRTVGVTHYLLGGIATTWAFFLARIIAVGMALRFPKFSQGLAQDPTTR     |
| 13ory | PATQPRALSIIQGRAVGVTHYLLGGIATTWAFFLARIIAVGMELRFPKFSQGLAQDPTTR    |
| 14pin | PAIQPRALSIVQGRAVGAAYHLLGGIVTTWAFFLARIIAVGMASRFPKFSQGLAQDPTTR    |
| 15por | PAIQPRALSITQGRAVGAAYHLLGGIGTTWAFFLARIISVGMATKFPKFSQALSQDPTTR    |
| 16syn | PAIQPRALSIIQGRAVGAAYHLLGGIVTTWAFFLARSLSIGMATKFPKFSQGLAQDPTTR    |
|       |                                                                 |
| 01can | RIWYGIATSHDFEADNVTEENLYQKIFASHFGHLAIIFLWTSGNLFHVWQGNFEAWIA      |
| 02chl | RLWFGIATAHDFESHGDMTEERLYQKIFASHFGQLAIIFLWTSGNLFHVWQGNFEQWVQ     |
| 03cya | RIWYGIATANDFETNDGITEENLYQKIFASHFGHLAIIFLWTSGNLFHVWQGNFEQWVK     |
| 04eug | RIWFGIATSHDFESHGDMTEENLYQKIFASHFGQLAIIFLWTSGNLFHVWQGNFEQWIK     |
| 05gui | RIWYGLATAHDFESHGDMTEENLYQKIFASHFGHLAIIFLWTSGNLFHVWQGNFEQWVL     |
| 06mai | RIWFGIATAHDFESHDDITEERLYQNIFASHFGQLAIIFLWTSGNLFHVWQGNFESWIQ     |
| 07mar | RIWFGIATAHDFESHDDITEERLYQKIFASHFGQLAIIFLWTSGNLFHVWQGNFEAWQ      |
| 08mes | RIWFGIATAHDFESHGDMTEENLYQKIFASHFGQLAIIFLWTSGNLFHVWQGNFERVWA     |
| 09nep | RIWYGIATAHDFESHGDMTEESLYQKIFASHFGQLAIIFLWTSGNLFHVWQGNFEQWQ      |
| 10nic | RIWFGIATAHDFESHDDITEERLYQNIFASHFGQLAIIFLWTSGNLFHVWQGNFESWVQ     |
| 11odo | RIWYGIATAHDLAHDGDMTEENLYQKIFASHFGHLAIIFLWTSAGNLFHVWQGNFEQWVA    |
| 12oen | RIWFGIATAHDFESHDDITEERLHQNIFASHFGQLAIIFLWTSGNLFHVWQGNFEAWVQ     |
| 13ory | RIWFGIATAHDFESHDDITEERLYQNIFASHFGQLAIIFLWTSGNLFHVWQGNFESWIQ     |
| 14pin | RIWFGIATAHDFESHDDITEERLYHKNIFASHFGQLAIIFLWTSGNLFHVWQGNFEAWVR    |
| 15por | RIWYGIATAHDFESHGDMTEENLYQKIFASHFGHLAIIFLWTSGNLFHVWQGNFEQWIL     |
| 16syn | RIWYGIATAHDFETHGDMTEENLYQKIFASHFGHLAIIFLWTSGLTFHVWQGNFEQWIK     |
|       |                                                                 |
| 01can | NPLKVKPIAHAIWDPHFGQAALKAFSRGGVNPVDSISYSGVYHWYTTIGMRTSSDLYAGA    |
| 02chl | DPLHIRPIAHAIWDPHFGQAAVEAFTRGASGPVNISTSGVYQWYTTIGIRTNQELYVGS     |
| 03cya | DPLNTRPIAHAIISDPHFGQRAIEAFSGAGSPVNIISYSGVYQWYTTIGMRTNEELYNGA    |
| 04eug | DPLHIRPIAHAIISDPHFGQPAIEAFSTRSQFPGPVNIIAYSGVYQWYTTIGLRTNVDLYNGS |
| 05gui | NPLKVKPIAHAIWDPHFGQPAVKAFTKGGVSPVNIISYSGVYHWYTTIGMRTNNDLYSGS    |
| 06mai | DPLHVRPIAHAIWDPHFGQPAVEAFTRGGAAGPVNIIAYSGVYQWYTTIGLRTNEDLYTGA   |
| 07mar | DPLHVRPIAHAIWDPHFGQPAVEAFTRGASGPVNIIAYSGVYQWYTTIGLRTNQDLYNGA    |
| 08mes | DPLHVRPIAHAIWDPHFGQPAVEAFTRGASGPVNIIAYSGVYQWYTTIGMRSNTDLYTGA    |
| 09nep | DPLHVRPIAHAIWDPHFGQPAVEAFTRGGAAGPVNIIAYSGVYQWYTTIGMRSNDLYTGA    |
| 10nic | DPLHVRPIAHAIWDPHFGQPAVEAFTRGALGPVNIIAYSGVYQWYTTIGLRTNEDLYTGA    |
| 11odo | KPLKVRPIAHSIWDPHFGESALKAFSGK-NTYPVNIIAYSGVYQWYTTIGFRTNQELYLGS   |
| 12oen | DPLHVRPIAHAIWDPHFGQPAVEAFTRGALGPVNIIAYSGVYQWYTTIGLRTNEDLYTGA    |
| 13ory | DPLHVRPIAHAIWDPHFGQPAVEAFTRGGAAGPVNIIAYSGVYQWYTTIGLRTNEDLYTGA   |
| 14pin | DPLHVRPIAHAIWDPHFGQPAIEAFTRGAPGPVNIIAYSGVYQWYTTIGLRTNEDLYAGA    |
| 15por | NPLKVKPIAHAIWDPHFGQPAKAFSGKGSAYPVNIIAYSGVYHWYTTIGMRTNQDLYTGA    |
| 16syn | DPLNIRPIAHAIWDPHFGGAVNAFTQAGASNPVNIIAYSGVYHWYTTIGMTNQELYSGA     |
|       |                                                                 |
| 01can | LFLVLALCMFAGRLHLQPKFKPSISWFKNNESRLNHHLSGLFGLSSLAWSGHLIHVAI      |
| 02chl | IFLLVLAGLFLFAGWLHLQPSFQPALSWFKNAESRLNHHLAGLFGVSSLAWTGHLVHVAI    |
| 03cya | IFLLILSALSIFAGWLHLQPKFRPNLSWFKNAESRLNHHLGGLFGTSSLAWTGHLVHVAI    |
| 04eug | MFLFLIATLALFAGWLHLQPKYSPKVSFWKDAESRLNHHLSALFGLSSLAWSGHLIHVAI    |
| 05gui | IFLLILASVMLFAGWLHLQPKFRPGLAWFKNNESRLNHHLSGLFGFSSVAWSGHLIHVAI    |
| 06mai | LFLFLSTLSLIGWLHLQPKWKPSLSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 07mar | LFLVLSSISLIGWLHLQPKWKPSLSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 08mes | LFLLLITASMTLIFAGWLHLQPKFKPSLSWFKNAESRLNHHLSGLFGVSSLAWTGHLIHVAI  |
| 09nep | LFLLVGAAILLIFAGWLHLQPKFKPSLSWFKNAESRLNHHLAGLFGVSSLAWTGHLVHVAI   |
| 10nic | LFLFLSAISLIGWLHLQPKWKPSVSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 11odo | VGLLLSCALLFAGWLHLQPKFRPSPSWFKNNESRLNHHLSGLMGVSSLAWTGHLVHVAI     |
| 12oen | LFLFLSAISLIGWLHLQPKWKPSVSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 13ory | LFLFLSTLSLIGWLHLQPKWKPSLSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 14pin | LFLFLSVIFLIFAGWLHLQPKFKPSLSWFKNAESRLNHHLSGLFGVSSLAWTGHLVHVAI    |
| 15por | LFLVLSAILLFGWLHLQPKFKPGLSWFKNNESRLNHHLSGLFGVSSLAWTGHLVHVAI      |
| 16syn | VFLVLASLFLFAGWLHLQPKFRPSPSWFKNAESRLNHHLAGLFGVSSLAWSGHLIHVAI     |



|       |                                                                 |
|-------|-----------------------------------------------------------------|
| 01can | P-ASRGQHIGWNNFLTTPPHAGLKPFFTGNWSDYSLNPDDPSHIFGTSSNSGKAILTFL     |
| 02chl | P-ESRGQHVGWDFLTVLPHAGLTPFFTGNWAAAYENPDLSQLFGTGGSGTAILTFL        |
| 03cya | P-ESRGQHVGWDFLQVAPHPAGLQPFFTGNWGVYTENPDANTHVFGSSDQAGTAILTFL     |
| 04eug | P-ESRGIHVRWDFLSQLPHPSGLEPFFKGNWSLYSENPDSLTHVFGTASGAGTAVLTFL     |
| 05gui | P-ESRGLHVGWDFNTKVYPHPEGLKAFFTGWWSNYAANPDADHIFGTTEGSGTAILTFL     |
| 06mai | PGSSRGGEYVRWNNFLDVLPPYQQLGPLLTGQWNLYAQNPSSSNHLFGTTQAGTAILTLL    |
| 07mar | P-ESRGEHVRWDFLTKLPHPEGLGPFFAGQWNLYAQNPVSSNHLFGTSQAGTAILTFI      |
| 08mes | P-ESRGQHVRWDFLNVLPHPAGLSPFFTGWAAAYAQNPDSSTSHIFSTSQAGTAILTFL     |
| 09nep | P-ESRGQHVGWDFLTTLPHAGLAPFFNGNWSVYAQNPDSASHLFGTSTGAGTAILTFL      |
| 10nic | P-ASRGEYVRWNNFLDVLPPHQLGPLFTGQWNLYAQNPSSSHLFGTAQAGTAILTLL       |
| 11odo | P-ASRGVHVWDFLTPPHAGLTPFFTGNWTVYAQNPSPSHVYGTSEGAGTRILTFL         |
| 12oen | P-GSRGESVRWNNFLDVLPHPEGLGPLFTGQWNLYAQNPSSSHLFGTSQSGTAILTLL      |
| 13ory | P-ASRGEYVRWNNFLDVLPPYQQLGPLLTGQWNLYAQNPSSSNHLFGTTQAGTAILTLL     |
| 14pin | P-ESRGVHVWDFLTVLPHPEGLEPLFTGQWNLYAQNPSSSHLFGTSQAGTAILTFL        |
| 15por | P-ESRGQHVGWDFNTTVLPHAGLQPFPSGNWSVYAQNPDTAQLFGTNEGAGTAILTFL      |
| 16syn | P-EARGQHVGWDFLTPPHAGLMPFFTGNWGVYAADPDATAGHIFGTSEGAGTAILTFL      |
|       |                                                                 |
| 01can | GGFHPQTQSLWLTIDIAHHHLAIAVIFIIAGHMYRTNFGIGHNIDKILEAHKPPSGKMLG    |
| 02chl | GGFHPQTQSLWLTIDMAHHHLAIAVVFILAGHMYRTIFGIGHSMREILEAQTPPSGRLGAG   |
| 03cya | GGFHPQTQSLWLTIDIAHHHLAIAVLFIVAGHMYRTNFGIGHSIKEILNGHRPPGGRLGAG   |
| 04eug | GGFHPETKSLWLTIDIAHHHLAIAVLFIVAGHMYRTNFAIGHRIDDLNAHKAAPSGKLG     |
| 05gui | GGFHPQTQSLWLTIDIAHHHLAIAGVIFIVAGHMYRTNWGIGHSLKEILDAHRAPSGRLGNG  |
| 06mai | GGFHPQTQSLWLTIDIAHHHLAIAFIFLIIAGHMYRTNFGIGHSIKDLLEAHTPPGRLGRG   |
| 07mar | GGFHPQTQSLWLTIDIAHHHLAIAVVFIIAGHMYRTNFGIGHSIKEILEHTPPGRLGRG     |
| 08mes | GGFHPQTQSLWLTIDIAHHHLAIAVLFIVAGHMYRTNFGIGHSMREILEAQRPPGGRLGAG   |
| 09nep | GGFHPQTQSLWLTIDMAHHHLAIAVVFILAGHMYRTNFGIGHSMREILEAHRAPSGRLGLG   |
| 10nic | GGFHPQTQSLWLTIDIAHHHLAIAFIFLIVAGHMYRTNFGIGHSMKDLDAHIPPGGRLGRG   |
| 11odo | GGFHPQTQSLWLTIDIAHHHLAIAVVFIIAGHMYRTNFGIGHNMKEILDAHRPPGGRLGAG   |
| 12oen | GGFHPQTQSLWLTIDMAHHHLAIAFIFLIIAGHMYRTNFGIGHSIKDLLEAHTPPGRLGRG   |
| 13ory | GGFHPQTQSLWLTIDIAHHHLAIAFIFLIIAGHMYRTNFGIGHSIKDLLEAHTPPGRLGRG   |
| 14pin | GGFHPQTQSLWLTIDMAHHHLAIALVFSIAGHMYRTNFGIGHSMREILEAHVPPGGLGRG    |
| 15por | GGFHPQSQSLWLTIDMAHHHLAIAVVFIVAGHMYRTNWGIGHNLDLDAHRPPSGRLGAG     |
| 16syn | GGFHPQTESLWLTIDIAHHHLAIAVIFIIAGHMYRTNWGIGHSIKEILNAHKGP--LTGAG   |
|       |                                                                 |
| 01can | HKGLFNTISNSLHFQLGLALACGLVLSLTAQHLYSIPPYAFISRDFVTQAALYTHHQYI     |
| 02chl | HKGLYDVTNNSLHFQLGLALASVGTICSLVAQHMYSLPPYAFLAQDFTTQASLYTHHQYI    |
| 03cya | HVGLYDVTNNSLHFQLGLALAGVITSLVAQHMYSLPPYAYLARDFTTQAALYTHHQYI      |
| 04eug | HFGLYETINNSLHFQLGLALASLGVITSLVAQHMYSLSPYAFILQDRTTMAALYTHHQYI    |
| 05gui | HKGLFETISNSLHFQLGLALASLGVITSLVAQHMYSLPSYAFIAKDYYVTQAALYTHHQYI   |
| 06mai | HKGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPAYAFIAQDFTTQAALYTHHQYI    |
| 07mar | HKGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPPYAFLAQDFTTQAALYTHHQYI    |
| 08mes | HSGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLSPYAFLAQDFTTQAALYTHHQYI    |
| 09nep | HKGLFDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPPYAFMAQDFTTMSALYTHHQYI    |
| 10nic | HKGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPAYAFIAQDFTTQAALYTHHQYI    |
| 11odo | HVGLFETITNSLHFQLGLALACGLVATSLTAQHMYALTPTYAYLSKDFTTEAALYTHHQYI   |
| 12oen | HKGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPAYAFIAQDFTTQAALYTHHQYI    |
| 13ory | HKGLYDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPSYAFIAQDFTTQAALYTHHQYI    |
| 14pin | HKGLYNTINNSLHFQLGLALASLGVITSLVAQHMYSLPPYAFIAQDFTTQAALYTHHQYI    |
| 15por | HRGLFDVTNNSLHMLQLGLALAGVITSLVAQHMYAMPYAFMAKDFTTQASLYTHHQYI      |
| 16syn | HTNLVDVTNNSLHFQLGLALASLGVITSLVAQHMYSLPSYAFIAQDHTTQAALYTHHQYI    |
|       |                                                                 |
| 01can | AGFLMVGAFAHGAIFFRVDYDPKQKNDVLYRMLEHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 02chl | AGFIMCGAFAHGAIFFRVDYDPEANRGNVLRVLDHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 03cya | AGFLMVGAFAHGAIFLVRDYDAEQKNNVLRARIIDHKEAIIISHLWSVSLFLGFHTLGLYV   |
| 04eug | AGFIMTGAFAHGAIFFIRDFDEEKNGNVLRIIDHKEAIIISHLWSVTLFLGFHTLGLYV     |
| 05gui | AGFLMVGAFAHGAIFFRVDYDPEQKNNVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 06mai | AGFIMTGAFAHGAIFFIRDYDPEQKNNVLRMLLEHKEAIIISHLWSVSLFLGFHTLGPYV    |
| 07mar | AGFIMTGAFAHGAIFFIRDYDPEQKNDVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 08mes | AGFIMTGAFAHGAIFFIRDYDPELNKNDVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV   |
| 09nep | AGFIMTGAFAHGAIFFIRDYDPIQNEGNVLRMLLEHKEAIIISHLWSVTLFLGFHTLGLYV   |
| 10nic | AGFIMTGAFAHGAIFFIRDYDPEQKNNVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 11odo | AGFLMVGAFAHGAIFFVRDYDPELNKNNVLRMLLEHKEAIIISHLWSVSLFLGFHVGLGYI   |
| 12oen | AGFIMTGAFAHGAIFFIRDYDPEQKNNVLRMLDHEAIIISHLWSVSLFLGFHTLGLYV      |
| 13ory | AGFIMTGAFAHGAIFFIRDYDPEQKNNVLRMLDHEAIIISHLWSVSLFLGFHTLGLYV      |
| 14pin | AGFIMTGAFAHGAIFLIRDYDPEQKNDVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV    |
| 15por | AGFLMVGAFAHGAIFFRVDYDPEQKNGNVLARMLEHKEAIIISHLWSVTLFLGFHTLGLYV   |
| 16syn | AGFLMVGAFAHGAIFFVRDYDPVANKNDVLRMLLEHKEAIIISHLWSVSLFLGFHTLGLYV   |
|       |                                                                 |
| 01can | HNDVVVAFGNPEKQILIEPFAQWIIQATSGKTYLGFNVLLASSSSSATQAA---QSLWLP    |
| 02chl | HNDVVQAFGTPEKQILIEPVFAQWIIQAAGKTYGDFDLSATSAPSLAG---QSLWLP       |
| 03cya | HNDVVQAFGTPEKQILIEPVFAQWIIQSVHGKSLYGFEVLLNNADISITRVPGSAQPIWLP   |
| 04eug | HNDVMQAFGTPEKQILIEPVFAQWIIQSAHGKNIYELNILLSDDSSNFAFSAS---QAIWLP  |
| 05gui | HNDVVVAFGTPEKQILIEPVFAQWIIQASSGKAMYGFVLLSANTSIAKNAS---SNIWLP    |
| 06mai | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKTYGDFDILLSSTNGTFNAG---RNIWLP     |
| 07mar | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKALYGFDVLLSSTNPNFANAG---QSIWLP    |
| 08mes | HNDVMQAFGTPEKQILIEPVFAQWIIQASHGKSLYGFDVLLSSSSSFAASAS---DSIWLP   |
| 09nep | HNDVMLAFGTPEKQILIEPVFAQWIIQASHGKALYGFDVLLSSADSPATASAS---QSIWLP  |
| 10nic | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKTSYGFVLLSSTSGFANAG---RSIWLP      |
| 11odo | HNDTVVAFGQPEKQILFEPLFAEYIIQAASGKAVYFNVLLSSSTNPATIAAG---NQVWLP   |
| 12oen | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKTSYGFVLLSSTNSPFAFNAG---RSIWLP    |
| 13ory | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKTYGDFDILLSSTSGFANAG---RTLWLP     |
| 14pin | HNDVMLAFGTPEKQILIEPFAQWIIQSAHGKTYGDFDILLSSTSGPSFDAG---KSIWLP    |
| 15por | HNDTMIAFGTPEKQILIEPVFAQWIIQASSGKALYGFDVLLSSSSNIATQAG---SNIWLP   |
| 16syn | HNDVVVAFGTPEKQILIEPVFAQWIIQATSGKALYGFDVLLSNPDSIASTTG---AA---WLP |

|       |                                                                |
|-------|----------------------------------------------------------------|
| 01can | NWLEAINNNNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY   |
| 02chl | GWLQGINSDTNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 03cya | GWLDAINSGNNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 04eug | GWLNGINDKSTSLFLQIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 05gui | GWLEAINSGKNSFLPIGPGDFLVHHAIALALHTTTTLILVKGALDARGSKLMPDKKDFGY   |
| 06mai | GWLNAVNNNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 07mar | GWLDAINNNNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY |
| 08mes | GWLDAINNSNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY |
| 09nep | GWLDAINSSNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY |
| 10nic | GWLNAVNNNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 11odo | GWLEAINNSKTDLFLKIGPGDFLVHHAIALGLHVTTLILVKGALDARGSKLMPDKKDFGY   |
| 12oen | GWLNAINNSNSNSFLFTIGPGDFLVHHAIALGSHTTTTLILVKGALDARGSKLMPDKKDLGY |
| 13ory | GWLNAVNNNSNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 14pin | GWLNAINDNNNSFLFTIGPGDFLVHHAIALGLHTTTTLILVKGALDARGSKLMPDKKDFGY  |
| 15por | GWLEAINSGKNSFLFTIGPGDFLVHHAIALGLHTTALILVKGALDARGSKLMPDKKDFGY   |
| 16syn | GWLDAINSGTNSFLFTIGPGDFLVHHAIALGLHTTALILKALDARGSKLMPDKKDFGY     |
|       |                                                                |
| 01can | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTIWIQGNASQFNESSTY   |
| 02chl | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHLGIWQGNVQFNESSTY    |
| 03cya | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHLGVWQGNVAQFNESSTY   |
| 04eug | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 05gui | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 06mai | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 07mar | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 08mes | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 09nep | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 10nic | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 11odo | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 12oen | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWATLYHWHKHTLWQGFVAQFNESSTY     |
| 13ory | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 14pin | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTLWQGNVQFNESSTY     |
| 15por | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTIGWVTFFYHWHKHTIWIQGNATQFNESSTY   |
| 16syn | SFPCDGPGRGGTCDISAWDAFYLVFWMLNTLGLWTFYHWHKHLGVWQGNVAQFNENSTY    |
|       |                                                                |
| 01can | LMGWFRDYLWLNSSPLINGYNPYGMNLSVWAWMFLFAGHLVWATGFMFLISWRGYWQELI   |
| 02chl | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLIYATGFMFLISWRGYWQELI   |
| 03cya | LMGWFRDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLIYATGFMFLISWRGYWQELI    |
| 04eug | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 05gui | LMGWL RDYLWLNSSPLINGYNPYGMNLSVWAWMFLFAGHLIYATGFMFLISWRGYWQELI  |
| 06mai | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 07mar | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 08mes | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLIYATGFMFLISWRGYWQELI   |
| 09nep | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 10nic | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 11odo | LMGWL RDYLWLNSSPLINGYNPYGMNLSVWAWMFLFAGHLIYATGFMFLISWRGYWQELI  |
| 12oen | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLIYATGFMFLISWRGYWQELI   |
| 13ory | LMGWL RDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI   |
| 14pin | LMGWSRDYLWLNSSQLINGYNPFGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI    |
| 15por | LMGWFRDYLWLNSSPLINGYNPYGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI    |
| 16syn | LMGWFRDYLWANSACLINGYNPYGMNLSVWAWMFLFGHLVWATGFMFLISWRGYWQELI    |
|       |                                                                |
| 01can | ESLVWAHERTPLANLITWKDKPVALSIVQARLVGLVHFSVGVYLTAAAFVIASTAGKFGM   |
| 02chl | ETLAWAHERTPLANLVRWRDKPVALSIVQARLVGLTHFSVGVYLTAAAFIASTSGKFGM    |
| 03cya | ETLVWAHERTPLANLVRWKDKPVALSIVQARLVGLAHFAVGYIFTYAAAFIASTASKFGM   |
| 04eug | ETLVWAHERTPITNVFKWTDKPVALSIVQARLVGLAHFSVGVYFTYAAAFIASTASKFGM   |
| 05gui | ETLVWAHERTPLANLVRWRDKPVALSIVQARLVGLVHFTVGYIFTYAAAFVIASTAGKFGM  |
| 06mai | ETLAWAHERTPLANLIRWRDKPVALSVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM    |
| 07mar | ETLAWAHERTPLANLVRWKDKPVALSIVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM   |
| 08mes | ETLAWAHERTPLANLVRWKDKPVALSIVQARVGLAHFSVGVYFTYAAAFIASTSGKFGM    |
| 09nep | ETLAWAHERTPLANLVRWKDKPVALSIVQARLVGLAHFSVGVYFTYAAAFVIASTSGKFGM  |
| 10nic | ETLAWAHERTPLANLIRWRDKPVALSIVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM   |
| 11odo | ETLVWAHERTPLANLIRWRDKPVALSIVQARLVGLVHFAVGYILTYAAAFVIASTSGKFAM  |
| 12oen | ETLAWAHERTPLANLIRWRDKPVALSIVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM   |
| 13ory | ETLAWAHERTPLANLIRWRDKPVALSIVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM   |
| 14pin | ETLAWAHERTPLANLVRWRDKPVALSIVQARLVGLAHFSVGYIFTYAAAFIASTSGKFGM   |
| 15por | ETLAWAHERTPLANLIRWKDKPVAMSIVQARLVGLAHFSVGVYLTAAAFVIASTAGKFGM   |
| 16syn | ETLVWAHERTPLANLVRWKDKPVALSIVQARLVGLAHFTVGVYLTAAAFIASTAGKFGM    |
|       |                                                                |
| 01can | VHVVKIYDTCIGCTQCVRACPDVLEMPWDGCKASQIASSPRTEDCIGCKRCETACPTD     |
| 02chl | SHTVKIYDTCIGCTQCVRACPTDLEMPWDGCKASQIASAPRTEDCVGCKRCESACPTD     |
| 03cya | AHTVKIYDTCIGCTQCVRACPTDLEMPWDGCRANQIASAPRTEDCVGCKRCESACPTD     |
| 04eug | SHSVKIYNTCIGCTQCVRACPTDLEMPWDGCKAGQIASSPRTEDCVCKRCESACPTD      |
| 05gui | SHSVKVYDTCIGCTQCVRACPDVLEMPWDGCKAGQIASAPRTEDCIGCKRCETACPTD     |
| 06mai | SHSVKIYDTCIGCTHCVRACPTDLEMPWDGCKAKQIASAPRTEDCVGCKRCESACPTD     |
| 07mar | AHAVKIYDTCIGCTQCVRACPTDLEMPWDGCKANQIASAPRTEDCVGCKRCESACPTD     |
| 08mes | AHSVKIYATCIGCTQCVRACPTDLEMPWDGCKANQIASAPRTEDCVGCKRCESACPTD     |
| 09nep | SHSVKIYDTCIGCTQCVRACPTDLEMPWDGCKAAQIASAPRTEDCVGCKRCESACPTD     |
| 10nic | SHSVKIYDTCIGCTQCVRACPTDLEMPWDGCKAKQIASAPRTEDCVGCKRCESACPTD     |
| 11odo | SHTVKIYDTCIGCTQCVRACPTDLEMPWDGCKSGQIASSPRVEDCVGCKRCETACPTD     |
| 12oen | SHSVKIYDTCIGCTQCVRACPTDLEMPWDGCKAKQIASAPRTEDCVGCKRCESACPTD     |
| 13ory | SHSVKIYDTCIGCTQCVRACPTDLEMPWDGCKAKQIASAPRTEDCVGCKRCESACPTD     |
| 14pin | AHSVKIYDTCIGCTQCVRACPTDLEMPWDGCKAKQIASAPRTEDCAGCKRCESACPTD     |
| 15por | AHSVKVYDTCIGCTQCVRACPDVLEMPWDGCKAKQIASAPRTEDCIGCKRCETACPTD     |
| 16syn | SHSVKIYDTCIGCTQCVRACPLDLEMPWDGCKAAQIASSPRTEDCVCKRCETACPTD      |

|       |                                                                |
|-------|----------------------------------------------------------------|
| 01can | FLSIRVYLG-AETSRSMGLTYMDKNLKYIYTPVISTLWLLTAGILIEINRFYPDALF      |
| 02chl | FLSVRVYLG-SETTRSMGLAYMK-DFTTYLSTAPVLTIVSLTAVAGLLIEINRFFPDALT   |
| 03cya | FLSIRVYLG-AETTRSMGLGY-M-SFSKYLSTAPVIGTLTAFFLAGLLIEINRFPDLLV    |
| 04eug | FLSVRVYLG-SETSRSMGLAYMK-YFTTYLSTAPVAVLWFTLTASLLIEINRFFPDIL-    |
| 05gui | FLSVRVYLG-GETTRSMGLAYMDNNFLKYLSTAPVLLTIWLSFTAALVIEANRFYPMPLY   |
| 06mai | FLSVRVYLG-PETTRSMGLSYM-R-DIKTYLSTAPVLTSLWFGALAGLLIEINRFPDALS   |
| 07mar | FLSVRVYLG-NETTRSMGLSYM-Q-DVKTYLSTAPVLATLWFGFLAGLLIEINRFFPDALV  |
| 08mes | FLSVRVYLG-NESTRSMGLAYMQ-DFQKYLSTAPVLATIWFIILAGLLIEINRFFPDALL   |
| 09nep | FLSVRVYLG-AETTRSMGLAYMK-DFTTYLSTAPVLAAVWFGFLAGLLIEINRFFPDALS   |
| 10nic | FLSVRVYLG-HETTRSMGLAYMR-DLKTYLSTAPVLTSLWFGALAGLLIEINRFFPDALT   |
| 11odo | FLSVRVYLGRALETTSLGLAYMN-DLQKYLSTAPVLLTLWMTFTAGFIIENRFFPDMLG    |
| 12oen | FLSVRVYLG-HETTRSMGLAYMR-DLKTYLSTAPVLSALWFGALAGLLIEINRFFPDALT   |
| 13ory | FLSVRVYLG-PETTRSMGLSYM-R-DIKTYLSTAPVVTSLWFGALAGLLIEINRFPDALS   |
| 14pin | FLSVRVYLG-HETTRSMGLAYMQ-DLKTYLSTAPVLAILCVSFLAALLIEINRFFPDALF   |
| 15por | FLSVRVYLG-AETTRSMGLAYMNNFTKYLSTAPVIGLWMTFTAGFIIENRFFPDVLY      |
| 16syn | FLSIRVYLG-AETTRSMGLAYMD-GLKSFLSTAPVMIMALLTFTAGILIEINRFPDLLF    |
|       |                                                                |
| 01can | YAF----MTVTLERRESTSLWERFCSWITSTENRLYIGWFGVLMIPCLLTATTVFIIAFI   |
| 02chl | AAF----MTAILERRESASLWARFCEWVTSTENRLYIGWFGVLMIPTLLTATSVFIIAFI   |
| 03cya | YPF----MTATLERNASVSLWEQFCGFTSTENRLYIGWFGVLMFPLLLTATTLFIIAFV    |
| 04eug | -----MISPVLLKYPAPSLWYRFCAWVASKKNRLYVWFGVLMIPTLLTAAITVFIIAFI    |
| 05gui | FPI----MTATLERRESASLWERFCSWITSTENRLYIGWFGVLMIPTLLTATTVFIIAFI   |
| 06mai | PPFF----MTAILERRESTSLWGRFCNWTSTENRLYIGWFGVLMIPTLLTATSVFIIAFI   |
| 07mar | LPFF----MTATLERRESASLWGRFCDWITSTENRLYIGWFGVLMIPTLLTATSVFIIAFI  |
| 08mes | VPMK----MTATLERRESANLWGRFCEFTSTENRLYIGWFGVLMIPCLLTATSVYIIAFV   |
| 09nep | FSFV----MTAILERRESTSVWARFCDWITSTENRLYIGWFGVLMIPTLLTATSVFIIIGFI |
| 10nic | FPFFSF--MTAILERRESESLWGRFCNWTSTENRLYIGWFGVLMIPTLLTATSVFIIAFI   |
| 11odo | LYF----MTATLERRREGVSLWERFCAWITSTENRLYIGWFGCLMFTLLTATSCYIIAFI   |
| 12oen | PPFFSF--MTAILERRESESLWGRFCNWTSTENRLYIGWFGVLMIPTLLTATSVFIIAFI   |
| 13ory | PPFFSF--MTAILERRESTSLWGRFCNWTSTENRLYIGWFGVLMIPTLLTATSVFIIAFI   |
| 14pin | LSLSFS--MTAILERRESANLWSRFCDWITSTENRLYIGWFGVLMIPTLLTATSVFIIAFI  |
| 15por | FYL----MTATLQRRESASLWERFCSWITSTENRLYIGWFGVLMIPTLLTATSVFIIAFV   |
| 16syn | HP-----MTTQLGLEQSLWSRFCCWITSTENRLYIGWFGVLMIPTLLTATTCFIIAFI     |
|       |                                                                |
| 01can | AAPPVDIDGIREPVSGSLLYGNIIITGAVVPTSNAIGLHLYPIWEAASLDEWLYNGGPYQ   |
| 02chl | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPTSNAIGLHFYPIWEAASLDEWLYNGGPYQ   |
| 03cya | AAPPVDIDGIREPVAGSLFYGNIIISGAVIPSSAAIGMHFYPWEAASLDEWLYNGGPYQ    |
| 04eug | AAPPVDIDGIREPVSGSLFYGNIIITGAVVPTSNAIGLHFYPIWEATSLDEWLYNGGPYQ   |
| 05gui | AAPPVDIDGIREPVAGSLLYGNIIITGAVIPSSASIGHFYPWEAASLDEWLYNGGPYQ     |
| 06mai | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPTSAAIGLHFYPIWEAASVDEWLYNGGPYE   |
| 07mar | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPTSAAIGLHFYPIWEAASVDEWLYNGGPYE   |
| 08mes | AAPPVDIDGIREPVSGSLLYGNIIISGVIPTMSNAIGLHFYPIWEAASLDEWLYNGGPYL   |
| 09nep | AAPPVDIDGIREPVSGSLFYGNIIISGAIIPSSAAIGHFYPIWEAASIDEWLYNGGCYE    |
| 10nic | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPSSAAIGLHFYPIWEAASVHEWLYNGGPYE   |
| 11odo | AAPPVDIDGIREPVAGSLLYGNIIISGAVIPSSNAIGMHFYPWEAASIDEWLYNGGPYQ    |
| 12oen | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPSSAAIGLHFYPIWEAASVDEWLYNGGPYE   |
| 13ory | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPSSAAIGLHFYPIWEAASVDEWLYNGGPYE   |
| 14pin | AAPPVDIDGIREPVSGSLLYGNIIISGAIIPSSAAIGLHFYPIWEAASVDEWLYNGGPYE   |
| 15por | AAPPVDIDGIREPVAGSLLYGNIIISGAVIPSSAAIGHFYPIWEAASLDEWLYNGGPYQ    |
| 16syn | AAPPVDIDGIREPIAGSLLYGNIIITAAVPPSSNAIGLHFYPIWEAASLDEWLYNGGPYQ   |
|       |                                                                |
| 01can | LVLHFLLGVAAYMGREWELSYRLGMRPWICVAFSAPVAAATAVFLIYPIGGQSFSDGMP    |
| 02chl | LIVCHFLLGICSYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 03cya | FVVMHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 04eug | LIVCHFLLGICSYMGREWELSFRLGMRPWIAVAYSAPVAAASAVFLIYPIGGQSFSDGMP   |
| 05gui | LIVDHFLLGVCYMGREWELSFRLGMRPWISVAFAPVAAASAVFLIYPIGGQSFSDGMP     |
| 06mai | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 07mar | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 08mes | MVCHFLLGIACYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP    |
| 09nep | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 10nic | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 11odo | LIVLHFLLGVAAYMGREWELSYRLGMRPWIFVAFSAPVAAASAVFLIYPIGGQSFSDGMP   |
| 12oen | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 13ory | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 14pin | LIVLHFLLGVAAYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFLIYPIGGQSFSDGMP   |
| 15por | LVLHFLTGVAAYMGREWELSYRLGMRPWISVAFAPVAAAAAVFLIYPIGGQSFSDGMP     |
| 16syn | LIVFHLIGIFCYLGRQWELSYRLGMRPWICVAFSAPVAAATATLLIYPIGGQSFSDGLP    |
|       |                                                                |
| 01can | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 02chl | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 03cya | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 04eug | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 05gui | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 06mai | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 07mar | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 08mes | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 09nep | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 10nic | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 11odo | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 12oen | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 13ory | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 14pin | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 15por | LGISGTFNFMLVFQAEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |
| 16syn | LGISGTFNFMLVLAQEHNLIMHPFHMAGVAGVGGSLFSAMHGSVLTSSLIRETTEENES    |

01can NYGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLGLWPVVGWLTSGISTM  
 02chl NEGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 03cya NAGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGLSTM  
 04eug NVGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGVSTM  
 05gui NYGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLGLWPVVGWLTALGIMTM  
 06mai NEGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 07mar NAGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 08mes NAGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTAMGISTM  
 09nep NAGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGVSTM  
 10nic NEGYRFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 11odo NYGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTAMGVSTM  
 12oen NEGYRFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 13ory NEGYRFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 14pin NAGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLAAWPVVGWLTALGISTM  
 15por NYAYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLGLWPVVGWLTALSVSTM  
 16syn NQGYKFGQEEETYNIVAAGGYFGRLLIFQYASFNNRSLHFFLGAWPVVGWIFAALAVCCF

01can AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGIEVMHERNAHNFPDLADNSLLPVAS  
 02chl AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDAV-----V  
 03cya AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLASGEVMPVAL  
 04eug AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----  
 05gui AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGIEVMHERNAHNFPDLASGESLPVAL  
 06mai AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----L  
 07mar AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----V  
 08mes AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----V  
 09nep AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----V  
 10nic AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----I  
 11odo AFNLNGLNFMNQSVVDSQGRVINTWADIINRADLGMEVMHERNAHNFPDLASGDVLPVAL  
 12oen AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----V  
 13ory AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----L  
 14pin AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLDA-----V  
 15por AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGMEVMHERNAHNFPDLASGESLPVAL  
 16syn AFNLNGLNFMNQSVVDSQGRVINTWADIINRANLGIEVMHERNAHNFPDLASGDAQMVAL

01can SSPSINSMALPWYRVHTVVLNDPGRLLISVHLMHTALVSGWAGSMALYELAVFDPSPDVLN  
 02chl EAPAVNGMGLPWYRVHTVVLNDPGRLLIAVHLMHTSLVSGWAGSMAPYELAVFDPSPDVLN  
 03cya TAPSIAMGLPWYRVHTVVLNDPGRLLIAVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 04eug -----MGLPWYRVHTVVLNDPGRLLISVHLMHTALVSGWAGSMALYELAFDPSPDVALN  
 05gui TAPAVIGMGLPWYRVHTVVLNDPGRLLIAVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 06mai EVPYLNGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVSGWAGSMALYELAVFDPSPDVLN  
 07mar EAPAVNGMGLPWYRVHTVVLNDPGRLLIAVHLMHTALVSGWAGSMALYELAVFDPSPDVLN  
 08mes EAPAVNGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVSGWAGSMAPYELAVFDPSPDVLN  
 09nep DAPAVQGMGLPWYRVHTVVLNDPGRLLIAVHLMHTALVSGWAGSMALYELAVFDPSPDVLN  
 10nic EAPSTNGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 11odo NAPAVNGMALPWYRVHTVVLNDPGRLLIAVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 12oen EAPSTNGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 13ory EVPSLNGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVSGWAGSMALYELAVFDPSPDVLN  
 14pin ESISIGGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 15por TAPAVNGMGLPWYRVHTVVLNDPGRLLIAVHLMHTALVAGWAGSMALYELAVFDPSPDVLN  
 16syn NAPAIEGMGLPWYRVHTVVLNDPGRLLISVHLMHTALVAGWAGSMALYELAFDSSDAVLN

01can PMWRQGMFVMPFMRIGVTSWGGWSITGESVSNPGLWSFEGVALTHIVLSGLLFLAAIW  
 02chl PMWRQGMFVLPFMRIGITSWGGWTSIGETAANPGVWSYEGVAAAHIVLSGLLFAAIIW  
 03cya PMWRQGMFVLPFMRIGITSWGGWTINGENVTPDGFWSFEGVAAAHIGLSGLLFLAAIW  
 04eug PMWRQGMFVLPFMRIGVTKSWGWSVTGESFSDPGIWSYEGVAAHIIISGLLFLAAIW  
 05gui PMWRQGMFVMPFMRIGVTSWGGWSITGESVSNPGFWSFEGVALAHIGLSGLLFLAAIW  
 06mai PMWRQGMFVLPFMRIGITSWGGWSISGGVTNPGIWSYEGVAAHIVFSGLCFLAAIW  
 07mar PMWRQGMFVLPFMRIGITSWGGWSITGETVTNAGIWSYEGVAAHIVLSGLLFLAAIW  
 08mes PMWRQGMFVLPFMRIGVTKSWGWDINGDSITDPLGWSYEGVAAHIIISGLLFLAAIW  
 09nep PMWRQGMFVLPFMRIGVTKSWGGWSITGESVSNPGIWSYEGVAAHIIISGLLFLAAIW  
 10nic PMWRQGMFVLPFMRIGITSWGGWSITGGVTNPGIWSYEGVAAHIVFSGLCFLAAIW  
 11odo PMWRQGMFVMPFMRIGITSWGGWSITGESVSNPGIWSFEGVALSHIISGMCFLLAAIW  
 12oen PMWRQGMFVLPFMRIGITSWGGWSITGGVTNPGIWSYEGVAAHIVFSGLCFLAAIW  
 13ory PMWRQGMFVLPFMRIGITSWGGWSISGGVTNPGIWSYEGVAAHIVFSGLCFLAAIW  
 14pin PMWRQGMFVLPFMRIGVTKSWGWNITGETVTNPGIWSYEGVAAHIVFSGLCFLAAIW  
 15por PMWRQGMFVMPFMRIGVTSWGGWSITGESVSNPGLWSLEGVALTHIVLSGLLFLAAIW  
 16syn PMWRQGMFVLPFMRIGVTSWGGWSITGETDLPDGFWSFEGVAAAHIVLSGLLFLAAIW

01can HWVYWDLELFRDPRTLEPALDLPKIFGIHLVLSLLCFGFGAFHVTGLFGPGIWSDAYG  
 02chl HWVYWDLELFRDPRTSNPALDLPKIFGIHLFLSGVLCFGFGAFHVTGLFGPGIWSDPYG  
 03cya HWVYWDLELFRDPRTGEPALDLPKMFGIHLFLSGLLCFGFGAFHVTGLFGPGMWSDAYS  
 04eug HWVYWDLELFRDPASGELKLDLPKIFGIHLFLSGALCLAFGVHVTGLFGPGIWSDPYG  
 05gui HWVYWDLELFRDPRTGNPALDLPKIFGIHLVLGLCFGFGAFHVTGAWGPGIWSDAYG  
 06mai HWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLAGVACFGFGAFHVTGLFGPGIWSDPYG  
 07mar HWVYWDLELFRDERTGKPSLDLPKIFGIHLFLSGVLCFAGFAGHVTGLFGPGIWSDPYG  
 08mes HWVYWDLELFRDPRTGEPALDLPKIFGIHLFLSGLLCFGFGAFHVTGLFGPGIWSDPYG  
 09nep HWVYWDLELFRDPRTGEPALDLPKIFGIHLFLSGLLCFGFGAFHVTGLFGPGIWSDPYG  
 10nic HWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLSGVACFGFGAFHVTGLFGPGIWSDPYG  
 11odo HWVYWDLELFRDPRTGEPALDLPKIFGIHLFLSGLLCFGFGAFHVTGLFGPGIWSDAYG  
 12oen HWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLSGVACFGFGAFHVTGLFGPGIWSDPYG  
 13ory HWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLAGVACFGFGAFHVTGLFGPGIWSDPYG  
 14pin HWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLSGVACFGFGAFHVTGLFGPGIWSDPYG  
 15por HWVYWDLELFRDPRTGEPALDLPKIFGIHLFLSGLLCFGFGAFHVTGLFGPGMWSDYG  
 16syn HWVYWDLELFRDPRTGESALDLPKMFGIHLFLSGLLCFGFGAFHVTGLFGPGMWSDPYG

01can LTGRIQSVAPAWGPEGFNPFNPGGIASHHIAAGTVGILAGVFHLLNVRPPQRLYRALRMGN  
 02chl ITGTQAVAPSWDATGFDPPYNPGGIASHHIAAGILGVLGAGLHLSVRPPQRLYNGLRMGN  
 03cya ITGRVQPVAPAWGPEGFNPFNPGGVSSHIAAGIVGILAGLHLSVRPPQRLYKALRMGN  
 04eug LSGKIEPVIPSWGAEGFDPYNVGGIASHHIAAGLLGLIAGGFHVLVRPSQRLFVLLRMGN  
 05gui ITGKVQPVAPTWGPEGFNPFNPGSVASHHIAAGILGFIAGIFHIAVRPPQRLYRALRMGN  
 06mai LTGKVQAVNPAPWGAEGFDPFVPGGIASHHIAAGTLGILAGLHLSVRPPQRLYKGLRMGN  
 07mar LTGKVQPVAPAWGAEGFDPFVPGGIASHHIAAGILGILAGLHLSVRPPQRLYKGLRMGN  
 08mes ITGRVQPIEPSWGADGFDPPNPGGIASHHIAAGILGILAGLHLSVRPSFRLYKALRMGN  
 09nep ITGSVQPVPEPAWGPEGFDPFNPGGIASHHIAAGILGILAGLHLSVRPPQRLYKALRMGN  
 10nic LTGKVQPVNPAPWVEGFDPFVPGGIASHHIAAGTLGILAGLHLSVRPPQRLYKGLRMGN  
 11odo VTGKVQPVAPAWGADGFNPFNPGGIAAHHIAAGIFGIFAGIFHLSVRPPQRLYRALRMGN  
 12oen LTGEVQPVNPAPWVEGFDPFVPGGIASHHIAAGTLGILAGLHLSVRPPQRLYKGLRMGN  
 13ory LTGKVQAVNPAPWGAEGFDPFVPGGIASHHIAAGTLGILAGLHLSVRPPQRLYKGLRMGN  
 14pin LTGKIQPVDPAPWGAEGFDPFVPGGIASHHIAAGILGILAGLHLSVRPPQRLYVGLRMGN  
 15por VTGKVLVPAPAWGPEGFNPFNPGGVASHHIAAGTVGILAGVFHLSVRPPQRLYRALRMGN  
 16syn LTGHVQPVAPWGPAGFNPFNPGGVAAHIIAAGIVGIIAGLHLSVRPPQRLYKALRMGN

01can IETVLSSSIIAAVFFASFVSGTMWYGAATPIELFGPTRYQWDSGYFQOEIEKRVEESLS  
 02chl IETVLSSSIIAAVFFAAVVSMTWYGSAAATPIELFGPTRYQWDLGFFQOEIERRVQTNLS  
 03cya IETVLSSSIIAAVFFAAIVAGTMWYGSAAATPIELFGPTRYQWQGYFQHEMERRVQKQDVA  
 04eug IETVLSSSIIAAVFFAAVVSMTWYGSAAATPIELFGPTRYQWQGYFQHEIERRVQASLS  
 05gui IETVLSSSIIAAVFFAAITVTGTMWYGSATTPIELFGPTRYQWDSGYFQOEIERRVENSLN  
 06mai IETVLSSSIIAAVFFAAVVSMTWYGSATTPIELFGPTRYQWQGYFQOEIYRRVSDGLA  
 07mar IETVLSSSIIAAVFFAAVVSMTWYGSAAATPIELFGPTRYQWQGYFQOEIDRRIRSSKA  
 08mes IETVLSSSIIAAVFFAAVVSMTWYGSAAATPIELFGPTRYQWDLGYNKEINRVRQASIA  
 09nep IETVLSSSIIAAVFFAAVVSMTWYGSAAATPIELFGPTRYQWQGYFQOEIEKRVRQSLA  
 10nic IETVLSSSIIAAVFFAAVVSMTWYGSATTPIELFGPTRYQWQGYFQOEIYRRVSAGLA  
 11odo IETVLSSSIIAAVFFAAVVSMTWYGAATPIELFGPTRYQWDSGYFQOEIERQVETSVS  
 12oen IETVLSSSIIAAVFFAAVVSMTWYGSATTPIELFGPTRYQWQGYFQOEIYRRVSGAGLA  
 13ory IETVLSSSIIAAVFFAAVVSMTWYGSATTPIELFGPTRYQWQGYFQOEIYRRVSDGLA  
 14pin IETVLSSSIIAAVFFAAIVAGTMWYGSATTPVELFGPTRYQWQGYFQOEIDRRVRAGLA  
 15por IETVLSSSIIAAVFFAAITVTGTMWYGSATTPIELFGPTRYQWDSGYFQOEIEKRVENAIA  
 16syn IETVLSSSIIAAVFFAAVVSMTWYGNATTPIELFGPTRYQWQGYFQOEIQRVDSQLA

01can NGLSLPEAWSNIPDKLAFYDYGNNPAKGGLFRAGPMNKGDIAGLWGHVPFQDKKEGHE  
 02chl EGKSASQAWAEIPEKLAFYDYGNNPAKGGLFRAGMNSGDGIAVGLGHAVFKEKQGNE  
 03cya AGASLSEAWNRIPAKLAFYDYGNNPAKGGLFRAGPMNKGDIAGLWGHATFQDKKEGRE  
 04eug DGCSLSEAWGAISPDLAFYDYGNNPAKGGLFRSGPMNKGDIATLWGLGHAVFIDKEGNS  
 05gui EGLSLSEAWSRIPDKLAFYDYGNNPAKGGLFRAGPMNKGDIAGLWGHVPFQDKKEGRE  
 06mai ENLSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNDGDIAGLWGHVPFRDKKEGRE  
 07mar ENLSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNDGDIAGLWGHAVFQDKKEGRE  
 08mes SGSTASEAWSRIPDKLAFYDYGNNPAKGGLFRAGMNSGDGIAAGLGHAVFQDKKEGRE  
 09nep SGASLSDAWAKIPEKLSFYDYGNNPAKGGLFRAGMNSGDGIAAGLGHVPFQDKKAGNE  
 10nic ENQSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNDGDIAGLWGHPIFRDKKEGRE  
 11odo EGLSESQAWSRIPDKLAFYDYGNNPAKGGLFRAGPMNKGDIAGLWGHPIFRDKDGRE  
 12oen KNQSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNSGDGIAVGLGHPIFRDKKEGRE  
 13ory ENLSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNDGDIAGLWGHPIFRDKKEGRE  
 14pin ENLSLSEAWSKIPEKLAFYDYGNNPAKGGLFRAGMNDGDIAGLWGHPIFRDKKEGRE  
 15por DGAAPSEAWSRIPDKLAFYDYGNNPAKGGLFRAGPMNKGDIAGLWGHVPFQDKKEGRE  
 16syn EGASLSEAWSTIPEKLAFYDYGNSPAKGGLFRAGMNSGDGIAQEWGHPFQDKKEGRE

01can LIVRRMPAFFENFPIILVDKDGIRADIPFRRAESKYSIEQVGVTCSFYGGKLNNQSFKD  
 02chl LFVRRMPTFFETFPVVLVDKDGVRADVPFRRAESKYSIEQVGVSFTFYGGELDGVTFND  
 03cya LTVRRMPTFFETFPVVLVDKDGVRADIPFRRAESKYSIEQMGVTVSFYGGKLDGQTFND  
 04eug LFVRRMPTFFETFPVILLDQNGVRADIPFRRAESKYSIEQVGVTFRFGGSFDTLSFND  
 05gui LTVRRMPTAFFETFPVILVDKDGIRADIPFRRAESKYSIEQVGVTVSFYGGKLNQGTFTD  
 06mai LFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYSIEQVGVTVEFYGGELNGVSYSND  
 07mar LFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYSIEQVGVTVEFYGGELDGVSFSD  
 08mes LFVRRMPTFFETFPVVLVDKDGIVRADIPFRRAESKYSIEQVGVSFAFYGGELDGVTFKD  
 09nep LFVRRMPTFFETFPVVLVDKDGVRADVPFRRAESKYSIEQVGVSFTFYGGELDGVTFND  
 10nic LFVRRMPTFFETFPVVLVDGDGIVRADVPFRRAESKYSIEQVGVTVEFYGGELNGVSYSND  
 11odo LTVRRMPTAFFETFPVILVDKDGIRADIPFRRAESKYSIEQVGVTDFYGGKLNQGTFTD  
 12oen LFVRRMPTFFETFPVVLVDGDGIVRADVPFRRAESKYSIEQVGVTIEFYGGELNGVSYSND  
 13ory LFVRRMPTFFETFPVVLVDEEGIVRADVPFRRAESKYSIEQVGVTVEFYGGELNGVSYSND  
 14pin LFVRRMPTFFETFPVVLVDKEGIVKADVPFRRAESKYSIEQVGVTVEFYGGELDRVSFGD  
 15por LSVRRMPAFFETFPVILVDKDGIRADIPFRRAESKYSIEQVGVTASFYGGKLNQGVFND  
 16syn LEVRRMPTFFETFPVIMTADGVVRADIPFRRAESKYSIEQVGVTVSFYGGALDQGTFSN

01can ASTVKKYARKAQGEVFEFDRITLSDGVFRSSPRGWFTFGHANFALLFFFGHLWHGSRT  
 02chl PATVKKYARRAQLGEIFEFDRTATLQSDGVFRASPRGWFTFAHLCFALLFFFGHIWHGART  
 03cya APTVKKYARKAQLGEAFEDRETATLQSDGVFRSSARGWFTFGHASFALLFFFGHIWHGART  
 04eug PATVKKYARHAQLGEIFDNRSITLQSDGVFRSSPRGWFTFGHLSFALLFFFGHIWHGART  
 05gui APTVKKYARKAQLGEVLEFDRITLSDGVFRSSPRGWFTFGHANFALLFFFGHLWHGSRT  
 06mai PATVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHATFALLFFFGHIWHGART  
 07mar PATVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHATFALLFFFGHIWHGART  
 08mes PTTVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHLSFALLFFFGHIWHGART  
 09nep PSTVKKYARRAQLGSVFEFDRTATLQSDGVFRSSPRGWFTFGHLSFALLFFFGHIWHGART  
 10nic PATVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHASFALLFFFGHIWHGART  
 11odo APTVKKFARKAQLGEVFEFDRITLSDGVFRSSPRGWFTFGHANFALLFFFGHLWHGART  
 12oen PATVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHASFALLFFFGHIWHGART  
 13ory PATVKKYARRSRLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHATFALLFFFGHIWHGART  
 14pin PAIVKKYARRAQLGEIFEFDRTATLQSDGVFRSSPRGWFTFGHATFALLFFFGHIWHGART  
 15por APSVKKYARKAQLGEVFEFDRITLSDGVFRSSPRGWFTFGHANFALLFFFGHLWHGSRT  
 16syn PSDVKKFARKAQLGEVFEFDRITLSDGVFRSPTSGWFTFGHASFALLFFFGHIWHGSRT

01can LFRDVFAGIGAETVEQVEFGVFQKVGDKTTKKQGYV-----ML  
02chl IFRDVFAGIDADLDEQVEFGAFKLGDSTTRRQSV-----MKNLYSLRRF  
03cya LFRDVFAGIGEETVEQVEFGAFQKVGDKTTRKQEA-----MKTLYSLRRF  
04eug LFKYLLAGIDPDLEEEIEFGTFEKLGDSTTKKELV-----  
05gui LFRDVFSGIGAETVEQVEFGAFQKLGDRSTKKQGAVMKVFLGWQLKINRMKTLVSLRRF  
06mai LFRDVFAGIDPDLDQVEFGTFQKVGDPTRRRQA-----MKILYSLRRF  
07mar LFRDVFAGIDPDLDQVEFGAFQKLGDPTRRRQVI-----MKILYSQRRF  
08mes IFRDVFAGIDPDLDQVEFGAFQKLGDASTRKQAV-----MKTLYSLRRF  
09nep IFRDVFAGIDPDLDQVEFGAFQKLGDPTRRRQAV-----MKNLYSLRRF  
10nic LFRDVFAGIDPDLDQVEFGAFQKLGDPTRRRQA-----MKTLYSLRRF  
11odo IFRDVFAGIGAETVEQVEFGAFQKLGDKSTKKQGA-----MKTLYSLRRF  
12oen LFRDVFAGIDPDLDQVEFGAFQKLGDPTRRRQAV-----MKTLYSLRRF  
13ory LFRDVFAGIDPDLDQVEFGAFQKLGDPTRRRQV-----MKILYSLRRF  
14pin LFRDVFAGIDSDLDRIEFGAFQKLGDPTRRRQV-----MKTLYSLRRF  
15por IFRDVFAGIGAETVEQVEFGAFQKLGDRSSKKQGAVMKVFLGWLLKINLMKTLVSLRRF  
16syn LFRDVFAGIDPDLEEQVEFGVFAKVGDLSTRKEA-----MKTLYSLRRF

01can FNKDK--NN---IGGRSIESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFTWGAMTLF  
02chl YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
03cya YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFTWGAMTLF  
04eug --METLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWTGAMNLF  
05gui YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWTGAMTLF  
06mai YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
07mar YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
08mes YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
09nep YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
10nic YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
11odo YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
12oen YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
13ory YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
14pin YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLF  
15por YHVETLFGSLVVGGRDQESTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWTGAMTLF  
16syn SPVVTLSNTS--MVGGRDLPGSTGFAWWSGNARLINLSGKLLGAHVAHAGLIVFWAGAMTLF

01can ETSHFIPEKPLYEQGMILLPHLATLWGVPAGGEIVNTYPYFATGVIHLVSSAVLFGGGI  
02chl EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
03cya EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
04eug EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
05gui EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
06mai EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
07mar EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
08mes EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
09nep EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
10nic EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
11odo EVSHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
12oen EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
13ory EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
14pin EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI  
15por EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGV  
16syn EVAHFVPEKPMYEQGLILLPHLATLWGVPAGGEIVNTYPYFVSGVLHLISSAVLFGGGI

01can YHSIVGPDVLEDSFSGYDWRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
02chl YHSLVGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
03cya YHAIIGPEVLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
04eug YHSLAGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
05gui YHSLIGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
06mai YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
07mar YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
08mes YHSLIGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
09nep YHSLIGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
10nic YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
11odo YHSLIGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
12oen YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
13ory YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
14pin YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
15por YHSLIGPDTLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA  
16syn YHALLGPETLEESFPFGYVWDRDKNKMTTILGIHLLGIGAFLLVIKALFIGGIYDTWA

01can PGGGDVRIITNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
02chl PGGGDVRIISNPTVSPGVIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
03cya PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
04eug PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
05gui PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
06mai PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
07mar PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
08mes PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
09nep PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
10nic PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
11odo PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
12oen PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
13ory PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
14pin PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
15por PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW  
16syn PGGGDVRIISNPTLNPATIFSYLLKSPFGGEGWIVGVNNMEDVIGGHIWIGVTCVIGGIW



01can HILTRPFSWARRAFVWSGEAYLSYSLGALALMGQTA AEYAWYNNNTVPSEFYGPTAAEAS  
 02chl HILTKPFAWARRAFVWSGEAYLSYSLGALALMGFTACCMSWFNTTAYPSEFYGPTGPEAS  
 03cya HILTKPFSWARRALVWSGEAYLSYSLGALALMGFIACCFVWFNNTAYPSEFFGPTGPEAS  
 04eug HICSRPSPWVRFTFVWSGEAYLSYSLGAVATMGFIAPMVWFNNTVPSEFYGPTGPEAS  
 05gui HILTKPFAWARRAFVWSGEAYLSYSLVAVSLMAFIASQYAWYNNNTVPSEFYGPTGPEAS  
 06mai HILTKPFAWARRAFVWSGEAYLSYSLGALSVFGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 07mar HILTKPFAWARRALVWSGEAYLSYSLGAIAVFGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 08mes HILTKPFAWARRAFVWSGEAYLSYSLGAIAVMGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 09nep HILTKPFWARRAFVWSGEAYLSYSLGAISVMGFIACCFVWFNNTVPSEFYGPTGPEAS  
 10nic HILTKPFAWARRALVWSGEAYLSYSLGALSVFGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 11odo HIFTKPFAWARRAFVWSGEAYLSYSLAAISLMGFTAALYSWYNNNTAYPSELYGPTGPEAS  
 12oen HILTKPFAWARRALVWSGEAYLSYSLAALSVFGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 13ory HILTKPFAWARRAFVWSGEAYLSYSLGALSVFGFIACCFVWFNNTAYPSEFYGPTGPEAS  
 14pin HILTKPFAWARRAFVWSGEAYLSYSLAALSFGFIACCFVWFNNTVPSEFYGPTGPEAS  
 15por HILTKPFAWARRAFVWSGEAYLSYSLGALSIMGLTASNFWYNNNTAYPSEFYGPTGPEAS  
 16syn HILTKPFGWARRALVWSGEAYLSYSLGALSIMGFIASVFWWFNNTAYPSEFYGPTGMEAS

01can QAQAFTLVRDQRLGANIASTQGPITGLGKYLMSRPTGEVILGGETMRFWDLRAPWLEPLR  
 02chl QSQTFTFLVRDQRLGANVASAQGPITGLGKYLMSRPTGEIIFGGETMRFWDFRGPWLEPLR  
 03cya QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDFRAPWLEPLR  
 04eug QAQAFTLVRDQRLGTNIAAQGPITGLGKYLMSRPTGEIIFGGETMRFWDFRGPWLEPLR  
 05gui QSQAFTFLVRDQRLGASVSSAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDLRAPWLEPLR  
 06mai QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEVIFGGETMRFWDLRAPWLEPLR  
 07mar QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEIIFGGETMRFWDLRAPWLEPLR  
 08mes QAQAFTLVRDQRLGANIASAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDCRAPWLEPLR  
 09nep QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEIIFGGETMRFWDLRAPWLEPLR  
 10nic QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEVIFGGETMRFWDLRAPWLEPLR  
 11odo QSQAFTFLVRDQRLGANVSSAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDLRAPWLEPLR  
 12oen QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEVIFGGETMRFWDLRAPWLEPLR  
 13ory QAQAFTLVRDQRLGANVGSAAQGPITGLGKYLMSRPTGEVIFGGETMRFWDLRAPWLEPLR  
 14pin QAQAFTLVRDQRLGASVGSAAQGPITGLGKYLMSRPTGEIIFGGETMRFWDLRAPWLEPLR  
 15por QAQAFTLVRDQRLGANVSSAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDLRAPWLEPLR  
 16syn QSQAFTFLVRDQRLGANIASAQGPITGLGKYLMSRSPSGEIIFGGETMRFWDFRGPWLEPLR

01can SSNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVATEINSVNYVSPRSWLTSSH  
 02chl GPNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 03cya GANGLDLNLKIKYDIQPWQERRAAEYMTAPLGSLSVGGVATEINSVNYVSPRSWLTSSH  
 04eug GPNGLDLNLKIKNDVQPWQERRAAEYMTAPLGSLSVGGVATEINAVNFVNRPSWLTSSH  
 05gui GPNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVATEINSVNYVSPRSWLTSSH  
 06mai GPNGLDLSRLKKDQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 07mar GPNGLDLSKLKKDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 08mes GPNGLDLNLKIKNDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNFVSPRSWLTSSH  
 09nep GPNGLDLSKLKNDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNFVSPRSWLTSSH  
 10nic GPNGLDLSRLKKDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 11odo GPNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVATEINSVNYVSPRSWLTSSH  
 12oen GPNGLDLSRLKKDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 13ory GPNGLDLSRLKKDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 14pin GPNGLDLSKLKKDIQPWQERRSAGEYMTAPLGSLSVGGVATEINAVNYVSPRSWLTSSH  
 15por GPNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVATEINSVNYVSPRSWLTSSH  
 16syn GPNGLDLNLKIKNDIQPWQERRAAEYMTAPLGSLSVGGVITDVNSVNYVSPRAWLTSSH

01can FFLGFFFIHGLWHAGRARAAGFEKGINRENEPVLMSRPLDMTIAIGRE----QERGW  
 02chl FCLGFFFFVGHWHAGRARAAGFEKGIIDRNEPVLMSRPLDMTIAIG---KTQEKRG  
 03cya FVLGFFFIHGLWHAGRARAAGFEKGLDRENEPVLMSKLLDMTIVAGRSD----SERGW  
 04eug FVLAFFFFVGHWHAGRARAAGFEKGIIDRSREIARKLPLDMTIFDLN---TENKNGW  
 05gui FFLGFFFIHGLWHAGRARAAGFEKGINRENEPVLTLRPIDMTIAGQE----EGRGW  
 06mai FVLGFFFFVGHWHAGRARAAGFEKGIIDRLPEPVYMTPLNMTIAGVR--VTKEENDL  
 07mar FVLGFFFFVGHWHAGRARAAGFEKGIIDRFEPVLSMTPLNMTIAGK--SSKEPKGL  
 08mes FTLAFFFFVGHWHAGRARAAGFEKGIERETEPVLFMKPLDMTISIDKS--VKKAEP  
 09nep FVLGFFFFVAHLWHAGRARAAGFEKGIIDRDEPTLFRPLDMTIAIG---TPEEKRG  
 10nic FVLGFFFFVGHWHAGRARAAGFEKGIIDRFEPVLSMTPLNMTIALGK--FTKDENDL  
 11odo FFLAFFFFIGHWHAGRARAAGFEKGINRANEPVLSMRPLDMTIAIGQN----QERGL  
 12oen FVLGFFFIHGLWHAGRARAAGFEKGIIDRFEPVLSMTPLNMTIALGK--FTKDEKDL  
 13ory FVLGFFFFVGHWHAGRARAAGFEKGIIDRLPEVLYMTPLNMTIALGR--VTKEENDL  
 14pin FVLGFFFFVGHWHAGRARAAGFEKGIIDRFEPVLSMTPLNMTIALGK--SSKEEKT  
 15por FFLGFFFIHGLWHAGRARAAGFEKGINRENEPVLMSRPLDMTIAIGQE----KTRGG  
 16syn FVLGFFFIHGLWHAGRARAAGFEKGIIDRETEPTLFRPLDMTIAVGRAP---VERGW

01can FDLDDWLKDRFVFIGWSGILLFPCAYLALGAWFTGTTFVSSWYTHGLASSYLEGCNFL  
 02chl FDVDDWLRRDRFVFIGWSGILLFPTAYLALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 03cya FDLDDWLKDRFVFIGWSGILLFPCAYLAVGAWFTGTTFTVSWYTHGLASSYLEGCNFL  
 04eug FDVADDWLKDRFVFIGWSGILLFPCAYLALGGWLTGTTFTVSWYTHGLASSYLEGCNFL  
 05gui FDLVDDWLKDRFVFIGWSGILLFPTSYSIGGWFTGTTFTVSWYTHGLASSYLEGCNFF  
 06mai FPLIDDLRRDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 07mar FDSMDWLRRDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 08mes FDLDDWLKDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNVL  
 09nep FDSMDWLRRDRFVFIGWSGILLFPCAYFAGGWLTGTTFTVSWYTHGLASSYLEGCNVL  
 10nic FDLDDWLRRDRFVFIGWSGILLFPCAYFAGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 11odo FDLIDDLKDRFVFIGWSGILLFPTAYLAAGGWMTGTTFTVSWYTHGLASSYLEGCNFL  
 12oen FDLIDDLKDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 13ory FDLIDDLKDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 14pin FDLVDDWLRRDRFVFIGWSGILLFPCAYFALGGWFTGTTFTVSWYTHGLASSYLEGCNFL  
 15por FDLVDDWLKDRFVFIGWSGILLFPCAYLAVGGWLTGTTFTVSWYTHGLASSYLEGCNFL  
 16syn FDLVDDWLKDRFVFIGWSGILLFPCAFMALGGWLTGTTFTVSWYTHGLASSYLEGANFL

01can TAAVSSPANSMGHSLFLWGPEAQDFTWCQIGGLWTFALHGSFGLIGFCLRQFEIAR  
 02chl TAAVSTPANSMGHSLFLWGPEAQDFTWCQIGGLWTFVALHGSFALIGFMLRQFEIAR  
 03cya TAAVSTPANSMGHSLFLWGPEAQDFTWCQIGGLWTFALHGAFLGIGFTLRQFEIAR  
 04eug TAAVSTPPNSMGHSLLLLLGSEAQWDFTRWLQIGGLWPFIALHGAFGLIGFMLRQFEIAR  
 05gui TAAVSTPANSMGHSLLLWGPEAQDFTWCQIGGLWAFALHGAFGLIGFCLRQFEIAR  
 06mai TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFALHGAFALIGFMLRQFEIAR  
 07mar TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFVALHGAFGLIGFMLRQFEIAR  
 08mes TAAVSTPANSMAHSLLLWGPEAQDFTWCQIGGLWTFALHGAFGLIGFMLRQFEIAR  
 09nep TAAVSTPANSMAHSLLLWGPEAQDFTWCQIGGLWTFIALHGSFGLIGFMLRQFEIAR  
 10nic TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFVALHGAFGLIGFMLRQFEIAR  
 11odo TAAVSTPANSMGHSLLLWGPEAQDFTWCQIGGLWAFIALHGAFGLIGFCLRQFEIAR  
 12oen TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFVALHGAFALIGFMLRQFEIAR  
 13ory TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFVALHGAFALIGFMLRQFEIAR  
 14pin TAAVSTPANSLAHSLLLWGPEAQDFTWCQIGGLWTFVALHGAFGLIGFMLRQFEIAR  
 15por TAAVSTPANSMGHSLFLWGPEAQDFTWCQIGGLWAFIALHGSFGLIGFCLRQFEIAR  
 16syn TVAVSSPADAFGHSLLFLWGPEAQGNLTRWFQIGGLWPFVALHGAFGLIGFMLRQFEISR

01can LVGLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRFLFLQGHNWTLN  
 02chl SVKIRPYNAIAFSAPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 03cya LIGLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRFLFFQGFHNWTLN  
 04eug AVQIRPYNAIAFSAPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 05gui LVGIRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSLGVAIIFRFLFFQGFHNWTLN  
 06mai PVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 07mar SVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 08mes AVKIRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 09nep AIGLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 10nic SVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 11odo LVGIRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRFLFLQGHNWTLN  
 12oen SVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 13ory SVQLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 14pin SVQLRPYNAIAFSAPIAVFVSVFLIYPLGQSGWFFAPSGVAAIFRILFFQGFHNWTLN  
 15por LVGLRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSLGVAIIFRFLFLQGHNWTLN  
 16syn LVGIRPYNAIAFSGPIAVFVSVFLIYPLGQSGWFFAPSGVAGIFRILFLQGHNWTLN

01can PFHMMGVAGILGALLCAIHGATVENTLFDGEDSNTFRAFTPTQSEETYSMTANRFS  
 02chl PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 03cya PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGNTFPAPNPQAEETYSMTANRFS  
 04eug PFHMMGVAGVLGAALLCAIHGATVENTIFEDGDSNTFRAFNPQSEETYSMTANRFS  
 05gui PFHMMGVAGILGALLCAIHGATVQNTIFEDGDAANTFRAFPTPTQSEETYSMTANRFS  
 06mai PFHMMGVAGVLGAVLLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 07mar PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQSEETYSMTANRFS  
 08mes PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDAANTFRAFNPQSEETYSMTANRFS  
 09nep PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 10nic PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 11odo PFHMMGVAGILGALLCAIHGATVENTLFDGEDAANTFRAFPTPTQSEETYSMTANRFS  
 12oen PFHMMGIVAGVLGAALLCAIHGATVEYTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 13ory PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 14pin PFHMMGVAGVLGAALLCAIHGATVENTLFDGEDGANTFRAFNPQAEETYSMTANRFS  
 15por PFHMMGVAGILGALLCAIHGATVQNTLFDGEDAANTFRAFPTPTQSEETYSMTANRFS  
 16syn PFHMMGVAGILGALLCAIHGATVENTLFDGEDSNTFRAFEPTQAEETYSMTANRFS

01can QIFGVAFANKRWLHFFLLFVPVTGLWSSIGIVGLALNLRAYDFVSQEIRAAEDPEFETF  
 02chl QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 03cya QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLGNLRAYDFVSQEIRAAEDPEFETF  
 04eug QIFGVAFSNKRWLHFFMVFPVTGLRMSALGIVGLALNLRAYDFVSQEIRAAEDPEFETF  
 05gui QVGVAFSNKRWLHFFMLFVPLAGLWTSAGIVGLALNLRAYDFVSQELRAEDPEFETF  
 06mai QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 07mar QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 08mes QIFGVAFANKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 09nep QIFGIAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQELRAEDPEFETF  
 10nic QIFGVAFSNKRWLHFFMLFVPVTGLWMSALGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 11odo QIFGVAFSNKRWLHFFMLFVPVTGLWTSAGIVGLALNLRAYDFVSQELRAEDPEFETF  
 12oen QIFGVAFSNKRWLHFFMLFVPVTGLWMSALGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 13ory QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPEFETF  
 14pin QIFGVAFSNKRWLHFFMLFVPVTGLWMSAIGVGLALNLRAYDFVSQEIRAAEDPESETF  
 15por QIFGVAFSNKRWLHFFMLFVPVTGLWTSAGIVGLALNLRAYDFVSQELRAEDPEFETF  
 16syn QIFGIAFSNKRWLHFFMLFVPVTGLWMSVIGVGLALNLRAYDFVSQELRAEDPEFETF

01can YTKNILLNEGIRAWMAAQDQPHENFVFPEEVLPRGNALMSGG-STGERPFSDIITSVRYW  
 02chl YTKNILLNEGIRAWMAAQDQPHKLVFPEEVLPRGNAL-MSG-ATGERPFSDILTSIRYW  
 03cya YTKNILLNEGIRAWMAAQDQPHENFVFPEEVLPRGNAL-MSG-TTGERPFSDIVTSIRYW  
 04eug YTKNILLNEGIRAWMAAQDQPHQFIFPEEVLPRGNAL-MAG-STGERPFSDIITSIRYW  
 05gui YTKNILLNEGIRAWMAAQDQPHENFIFPEEVLPRGNALMSGG-STGERPFSDIITSIRYW  
 06mai YTKNILLNEGIRAWMAAQDQPHENLIFPEEVLPRGNAL-MSG-STGERSFADIITSIRYW  
 07mar YTKNILLNEGIRAWMAAQDQPHENLVFPEEVLPRGNAL-MSG-NTGERPFADIITSIRYW  
 08mes YTKNILLNEGIRAWMAAQDQPHENLVFPEEVLPRGNAL-MAG-STGERPFSDIITSIRYW  
 09nep YTKNILLNEGIRAWMAAQDQPHKLVFPEEVLPRGNAL-MSG-STGERPFSDILTSIRYW  
 10nic YTKNILLNEGIRAWMAAQDQPHENLIFPEEVLPRGNAL-MSG-STGERSFADIITSIRYW  
 11odo YTKNILLNEGIRAWMAAQDQPHENFVFPEEVLPRGNALMSGG-STGERPFSDIITSVRYW  
 12oen YTKNILLNEGIRAWMAAQDQPHENLIFPEEVLPRGNAL-MSG-STGERSFADIITSIRYW  
 13ory YTKNILLNEGIRAWMAAQDQPHENLIFPEEVLPRGNAL-MSG-STGERSFADIITSIRYW  
 14pin YTKNILLNEGIRAWMAAQDQPHENLIFPEEVLPRGNAL-MSG-NTGERSFADIITSIRYW  
 15por YTKNILLNEGIRAWMAAQDQPHENFIFPEEVLPRGNALMSGG-STGERPFSDIITSVRYW  
 16syn YTKNILLNEGIRAWMAAQDQPHENFIFPEEVLPRGNAL-MSG-TTGERPFSDIVTSIRYW



```

01can VIHSITIPSLFIAGWLFVSTGLAYDVFGTTPRPNEYFSENRRQVPLINDRFNAREELDDL
02chl VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTEDRQETPLITDRFNALEQVKKFS
03cya VIHTVTIPSLFVAGWLFVSTGLAYDVFGTTPRPDEYFTESRQEIPLITDRFSTN-----
04eug VIHSVTIPSLFVGGWIFVSTGVVDYIFGTTPRPSEYFTETRQQAPLISDRFNALEQMDQFT
05gui IHSITIPALFVAGWLFVSTGLAYDVFGTTPRPNEYFTQERQQVPLVNDRFSAKQLEDLT
06mai VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTESRQEIPLITDRFDSLEQLDEFS
07mar VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTENRQEVPLITGRFNSLEQIDEFT
08mes VIHSITIPSLFVSGWLFVSTGLAYDVFGTTPRPNEYFTEDRQDIPLITDRFNALEQLNQYT
09nep VIHSITIPSLFVAGWLFVSTGLAYDVFGSPRPNEYFTESRQETPLITDRFNALEQMDILT
10nic VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTESRQEIPLITGRFDPLEQLDEFS
11odo IHSITIPSLFVSGWLFVSTGLAYDVFGTTPRPNEYFTQDRQIPLVNDRFSAKQLEDLT
12oen VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTESRQEIPLITGRFDSLEQLDEFS
13ory VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTESRQEIPLITDRFDSLEQLDEFS
14pin VIHSITIPSLFIAGWLFVSTGLAYDVFGSPRPNEYFTESRQEVPLVTRGRFDPLEQLDEFT
15por VIHSITIPALFVAGWLFVSTGLAYDVFGTTPRPNEYFTQDRQQVPLVNDRFNAKQLEDLT
16syn VIHSITIPMLFIAGWLFVSTGLAYDAFGTTPRPDEYFTQTRQELPILQERYDINQIEQEFN

01can SNLMAN--KPLQPISYPIFTFRWLAIHGLAIPTVFFFGAITAMQFIQR-----
02chl EVNMTARKS---YTYPIFTVRWLAVHALAVPTVFFLGAITAMQFIQRMATGTSKVKVSD
03cya ---MNN--PNQPVSYPIFTVRWLAIHGAIPAVFVIGSITAMQFIQR-----
04eug K--MTNKKD---TRYPIFTVRWLAVHALAIPTVFFLGSISAMQFIQRMAT--ISKNKTSN
05gui KGLMSK--IKQPISYPIFTFRWLAIHGLAVPTVFFLGAITSMQFIQR-----
06mai RSFMTIDR----TYPIFTVRWLAVHGLAVPTVFFLGSISAMQFIQRMAT--QTVEDSSR
07mar KSFMTIDR----TYPIFTVRWLAVHGLAVPTVFFLGAISAMQFIQRMAT--QIIDTTPK
08mes K--MAAEKT---IQYPIFTVRWLAVHALAVPTVFFLGAISAMQFIQR-----MADT
09nep EGLMTTNTQ---YIYPIFTVRWLAVHALAVPTVFFLGSISAMQFIQRMATGTSKAKADE
10nic RSFMTIDR----TYPIFTVRWLAVHGLAVPTVFFLGSISAMQFIQRMAT--QTVENSSR
11odo KGLMTK--NINQPVAIYPIFTFRWLAIHGLAIPTVFFLGGITAMQFIQR-----
12oen RSFMTIDR----TYPIFTVRWLAVHGLAVPTVFFLGSISAMQFIQRMAT--QTAESSR
13ory RSFMTIDR----TYPIFTVRWLAVHGLAVPTVFFLGSISAMQFIQRMAT--QTVEDSSR
14pin RSFMTIDR----TYPIFTVRWLAVHGLAIPTVFFLGSISAMQFIQRMAT--QTIDDTSK
15por KGIMTKGNANQPVSYPIFTFRWLAIHGLAIPTVFFLGAITSMQFIQR-----
16syn Q--MATQNPQNQVITYPIFTVRWLAVHGLAVPTVFFLGAITAMQFIQR-----

01can -MALKTRLGELLRLPLNSEYKGVAPGWGTTPLMGIFMALFLLIILQIYNSSLILENLD
02chl -TGVSTPLGTLKPLNSEYKGVAPGWGTTPLMGIFMALFAVFLVILQIYNSSVLLDDVT
03cya -MPQRTALGNILRLPLNSEYKGVAPGWGTTPLMGVFMFLFFVFLIIILQIYNSSLLENVQ
04eug SKGKTTTLGTLKPLNSEYKGVAPGWGTTAGIMLIFMTLFAIFLTIILEIYNSSVILD-IK
05gui -MALRTRLGELLRLPLNSEYKGVAPGWGTTPLMGVFMFLFFVFLIIILQIYNSSLILENVD
06mai PKPKRTGAGSLKPLNSEYKGVAPGWGTTPLMGVAMALFAIFLSIILEIYNSSVLLDGIL
07mar TKGKSGIGDILKPLNSEYKGVAPGWGTTPLMGIMMALFAVFLVILQIYNSSVLLDGVS
08mes SQGKRTVVGDFLKPLNSEYKGVAPGWGTTPLMGVFMALFAVFLVILQIYNASVLLDGIP
09nep STSKITPLGTALKPLNSEYKGVAPGWGTTPLMGVFMALFAVFLIILEIYNSSVLLDDVG
10nic SGPRTAVGDLLKPLNSEYKGVAPGWGTTPLMGVAMALFAVFLSIILEIYNSSVLLDGIS
11odo -MALRTRLGEILRLPLNSEYKGVAPGWGTTPLMGVAMALFAVFLIILEIYNSSVLIENVD
12oen ARPKKTGLGGLKPLNSEYKGVAPGWGTTPLMGVAMALFAVFLSIILEIYNSSVLLDGIS
13ory PGPRTQTRVGNLLKPLNSEYKGVAPGWGTTPLMGVAMALFAVFLSIILEIYNSSVLLDGIL
14pin TTPKETLVGTTLKPLNSEYKGVAPGWGTTPLMGVAMALFAVFLSIILEIYNSSVLLDGIP
15por -MALRTRLGEILRLPLNSEYKGVAPGWGTTPLMGVFMFLFFVFLIIILQIYNSSVLENVD
16syn -MAQRTRLGDILRLPLNSEYKGVVPGWGTTPVMGVFMALFLVFLIILEIYNSSVLIILEGFS

01can ISWTTLGI-MFTLKI FVYSCVAFFCSLFIFGFLSNDPSRNPGRKDL-----
02chl MSWESLSK-MTLKIFVYTVVTFVSLFIFGFLSNDPGRNPGQRELD-----
03cya VSWTAATA-MTLKIFVYTVVTFVSLFIFGFLSNDPARKPSRKDIR-----
04eug -----MLILKVFYALILIFVSLFVFGLLSNDPGRNPYDDTGELMREVFYKDPYKG
05gui VDWASLGN-MFTLKI FVYTVVTFVSLFIFGFLSNDASRNPGRKDL-----
06mai TN-----MLTLKLFVYTVVTFVSLFIFGFLSNDPGRNPGREE-----
07mar VSW-----MLTLKLFVYTVVTFVSLFVFGFLSNDPGRNPGRKE-----
08mes ANWSSLKMYMLTLKIFVYTVVTFVSLFIFGFLSNDTGRNPKRKDIE-----
09nep VSWYSLGK-MTLKIFVYTVVTFVSLFIFGFLSNDPGRNPGTKDF-----
10nic MN-----MLTLKLFVYTVVTFVSLFIFGFLSNDPGRNPGREE-----
11odo VDSWN-GIVMLTLKILVYTVVTFVSLFIFGFLSNDPSRNPGRKDL-----
12oen MN-----MLTLKLFVYTVVTFVSLFIFGFLSNDPGRNPGREE-----
13ory MN-----MLTLKLFVYTVVTFVSLFIFGFLSNDPGRNPGRDE-----
14pin VSWG-----MLTLKLFVYAVVVFISLIFGFLSNDPGRNPGRKE-----
15por VDWATLGS-MFTLKI FVYTVVTFVSLFIFGFLSNDPSRNPGRKDL-----
16syn VDWAG-----MLTKIAVYIVVGLFISLIFGFLSNDPGRNPGRKE-----

01can ----MNSTGRIPLWLVTGFGGIVVLTVLGIFIGYSYGLGSSL-----
02chl --MSNTGTTGRIPLWLVTGAGTAALTLVAVFFYGSYVGLGSSL-----
03cya ---MANTGRIPLWLVTAGLAAIGVLGIFFGYSYGLGSSSI-----
04eug PLMSNSENTGRIPLWLVTIIGLAAIALLALFFYGSYGLGSSL----MSFITLNNFFN
05gui ----MASTGRIPLWIIATFGGIAALTIVGLFIYGSYSGIGSAL-----
06mai ----MADTTGRIPLWLIGTVGILVIGLIGFFYGSYGLGSSLMNLSLTCICFNSVL
07mar ---MANTTGRVPLWLIGTVAGILVIGLIGFFYGSYGLGSSL----MFNIYLENAFY
08mes --MSNTGTTGRIPLWLVTGVLGAILGLLALFFYGSYGLGSSL-----MNIG
09nep --MSNTGTTGRVPLWFGMIVGLAALGLLIGFFYGSYGLGSSL-----MST
10nic ---MADTTGRIPLWIIIGTVAGILVIGLIGFFYGSYGLGSSLMNTFSLIGICLNSTL
11odo ----MSNTGRIPLWLVLVGLGAVITMLSLFYGAYSGLGSSL-----
12oen ---MADTTGRIPLWIIIGTVAGILVIGLIGFFYGSYGLGSSLMNLSLTCICFNSVI
13ory ----MADTTGRIPLWLIGTVGIAVIGLIGFFYGSYGLGSSLMNLSLTCICFNSVI
14pin ---MADTTGRIPLWLIGTVGIVIGLIGFFYGSYGLGSSL----MPVMFNIFLDDAFI
15por ----MAGTGRIPWLVLVATVGGMAITVLGIFIGYSYGLGSSL-----
16syn ---MFAEG-RIPLWVVGWVAGIAGVGLFFYGAYAGLGSSM-----

```

01can ---MINQLPETYNIFAPIIDIMPVPIILFLLAFVWQAAVGR--MSGINPNKQPVELN  
02chl ---VLLLAKLPEAYAPFDPIVDVLPVIVPLFLLALVWQASVSFR--MAKNPNKQSVELN  
03cya MLMSLF LAKLPAAYALFDPIVDILPIIPLFLLAFVWQAAIGFK--MVSQNPKNKQSVELN  
04eug ---TNTFFGOLLPEQYAPFDPLIDVLPPIIPLFLLAFVWQASVKFR--MAQTNPNKKTVELN  
05gui MEGILFLAKLPEAYAIFKPIIDVAPVPIPVFLLAFVWQAAVGR--MSGPNPNKQPVELN  
06mai CPTSFFFAKLPEAYAIFNPIDVMPVPIPLFLLAFVWQAAVSFR--MTQSNPNKQSVELN  
07mar --LNGITFAKLPEAYSIFDPIVDVMPPIIPLFLLAFVWQASVSFR--MTQPNPNKQSVELN  
08mes --FDMILAKLPEAYAMFDPLVDVLPVPIPLFLLAFVWQASVSFR--MTEPNPNKQSVELN  
09nep --LPILLATLPEAYLPRPLVDVLPVPIPLFLLAFVWQAAVSFR--MTKPNPNKQTVELN  
10nic FSSSFFFGKLPEAYAFNPIVDIMPVPIPLFLLAFVWQAAVSFRMTTQSNPNKQSVELN  
11odo --MESLLARLPEAYVVFSPIDVLPPIIPLFLLAFVWQAAIGFR--MTGPNPNKQSVELN  
12oen YSSRFFLAKLPEAYAFNPIVDVMPVPIPLFLLAFVWQAAVSFR--MTQSNPNKQSVELN  
13ory YPTSFFFAKLPEAYAIFNPIDVMPVPIPLFLLAFVWQAAVSFR--MTQSNPNKQSVELN  
14pin HSNPNPFGKLPEAYAISDPIVDVMPPIIPLFLLAFVWQAAVSFR--MTRPNPNKQSVELN  
15por MNSALFLAKLPEAYAVFKPIVDILPVIPLFLLAFVWQAAIGFR--MSGPNPNKQSVELN  
16syn METIYLLAKLPEAYQIFDPLVDVLPVPIPLFLLAFVWQAAVGFKNDRNPNPNKQSVELN

01can RTSLFWGLLLIFVLAVLFSSYFFN--MEK--ALILSIFISLGLITTSYSIYISFGPLSKTL  
02chl RTSLYWGLLLIFVLAVLFSSYIFN--MDSPAFFFTIFLWCLLSITGSIYVGFPPSQQL  
03cya RTSLFWGLLLIFVLAILFSSYIFN--MEI--STFLSIFISAALLGITGSIYTAFGPPSKNL  
04eug RTSLYWGLLLIFVLAVLFSSYIFNMFKIISFLSTIFLGFIIISTTIYSIYMGFGPMRKL  
05gui RTSLFWGLLLIFVLAVLFSSYFFN--MET--ATVLIIFIASLLGLTGSIYTAFGPNKSKEL  
06mai RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATLVAISISGLLVSTGYALYTAFGQPSQQL  
07mar RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATVAIFISCLLSIFGTGYALYTAFGQPSNEL  
08mes RTSLYWGLLLIFVLAILFSSYIFN--MET--ATIFSIFSCLLIGLTSYLSYTFSGNASSEL  
09nep RTSLYWGLLLIFVLAILFSSYIFN--METPALVATVFSCLVLSITGSIYIGFGPPSKEL  
10nic RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATLVAIFISGLLVSTGYALYTAFGQPSQQL  
11odo RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATIIVIFVSSLLGITAYSITTAFGPAKKNL  
12oen RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATLVAISISGLLVSTGYALYTAFGQPSQQL  
13ory RTSLYWGLLLIFVLAVLFSSYFFN--MET--ATLVAISISGLLVSTGYALYTAFGQPSQQL  
14pin RTSLYWGLLLIFVLAVFSSYFFN--MET--ATLVTISISCLLVSTGYALYTAFGQPSQQL  
15por RTSLFWGLLLIFVLAVLFSSYFFN--MET--ATVLSIFISLGLITGSIYTAFGPASKDL  
16syn RTSLYWGLLLIFVLAVLFSSYFFN--MES--ATVLSITFAVILIAITGLAVTSGFPPSAEL

01can RDPFEEHEEALVYVFLIGTLVVIFFAIFRDPPIAKK---MISVHTRKVIDNSGG  
02chl RDPFEEHEDMEALVYTFLLVGLTGIIFFAIFFREPPRIK---MIQPQTYLRVADNTGA  
03cya RDPFEEHEDMEALVYTFLLVTLGLIFFSIIFRDPPIRIN---MIQPQSYLTAADNSGA  
04eug KDPFEEHENMEALVYTFLLIGTLGVIFFAIFFREPPRIK---MIKPQTYLKIADNTGA  
05gui RDPFDDHEDMETVYTFLLIGTLAVLFAAVFRDPPIAKK---MIQTQTYLTVADNSGA  
06mai RDPFEEHGDMEALVYTFLLVSTGLIIFFAIFFREPPKVPKTK---MIQPQTYLTVADNSGA  
07mar RDPFEEHEDMEALVYTFLLVGLTGIIFFAIFFREPPKVPKTKMIQPQTYLTVADNSGA  
08mes RDPFEEHEDMEALVYTFLLVGLTGIIFFAIFFREPPRIK---MIQAQSYLTVADNSGA  
09nep RDPFDEHEDMEALVYTFLLISTGLIFFGIFFREPPRIK---MIQPQTYLTVADNSGA  
10nic RDPFEEHGDMEALVYTFLLVSTGLIIFFAIFFREPPKVPKTKN---MIQPQTYLTVADNSGA  
11odo RDPFEEHEDMEALVYTFLLIGTLVIFFAIFFREPPRIK---MIYPTQTYLTVADNTGA  
12oen RDPFEEHGDMEALVYTFLLVSTGLIIFFAIFFREPPKIKTKTKMIQPQTYLTVADNSGA  
13ory RDPFEEHGDMEALVYTFLLVSTGLIIFFAIFFREPPKVPKTKKMIQPQTYLTVADNSGA  
14pin RDPFDEHEDMEALVYTFLLVSTGLIIFFAIFFREPPKLPKGGKMIQSYTYLTVADNSGA  
15por RDPFEEHEEALVYVFLIGTLVVIFFAIFFREPPRIAK---MIQTQSYLTVADNSGA  
16syn GDFPDDHEDMESVAILVLTMAVLFFAIAIFREPPRIEK---MIQPQTYLTVADNSGA

01can KEIMCIRVLG-T-NRKYATLGDVIGVVKNAVNPMPVKKSDVVRAVVVRKKTINRKNFF  
02chl RELMCIRVLGGG-KQRAATVGDIIIAVVKDATPNMPVKRSDIVRAVIVRTKKNVRVNGT  
03cya KKLKMCIRVLGGG-NRRYARIGDVIAVVKDGINPIKKSDDTVKAVIVRTKELKRDNGM  
04eug KQIMCIRILG-P-NCQYANIGDIIIAVVKNAIPNMVVKKSDIVKAVIVRTKGVRESGM  
05gui KKMICIRILG-G-NRKYASIGDVIGVVKDATPNMPVKRSDIVRAVIVRTKKTIRKRDGM  
06mai RKLKMCIRVIGAAGNQRYARIGDVIIAVIKDAPVQMPLESEVIRAVIVRTKEFKGDDGI  
07mar RKLKMCIRVLGTS-NRKYANIGDIIIAVVKNAIPNMPIKKSSEIVRAVIVRTKEFKRNGM  
08mes KKMICIRVLGGS-QRKYAAGDVIGVVKDAPVNMVVKKSEVIRAVIVRTCKEIRRENGM  
09nep RKLKMCIRVLG--RKYASIGDVIIIAVVKDAIPNMVVKKSDVVRAVVVRTSKPVRDGTGM  
10nic RELMCIRIIGAS-NRRYAHIGDVIAVVKNAIPNMVVKKSEVIRAVIVRTCKELKRDNGM  
11odo RKLKMCIRVLG-G-NRKYAKIGDIIIAVVKNAIPNMPIKKSSEIVRAVIVRTSKTIRRPDGM  
12oen RELMCIRIIGAS-NRRYAHIGDIIIAVVKNAIPNMVVKKSEVIRAVIVRTCKELKDDGI  
13ory RKLKMCIRVIGAASNQRYARIGDVIAVVKDAPVQMPLESEVIRAVIVRTCKEFKCEDGI  
14pin RKLKMCIRVLGAS-NRKYAHIGDVIIIAVVKNAIPNMVVKKSEVIRAVIVRTCKEFDNGM  
15por RKLKMCIRVLGTS-NPSYASIGDVIGVVKDAPVNMVVKKSDVVRAVVVRTKALRRDNGM  
16syn RKLKMCIRVLGTS-NCTYGGIGDIIIAVVKDAPVNMVVKKSDVVRAVIVRTKQPLRRASGM

01can SVRFDDNAAVIGSDG-NPKGTRVFGPVARELDRDNFSKIASLAQEVVLMSPKTKYRK  
02chl SIRFDENAAVINKEN-NPRGTRVFGPVARELDRDNFTKISLAQEVVLMSPKTKYRK  
03cya NICFDNAAVINKEN-NPRGTRVFGPVARELDRDNFTKISLAQEVVLMSPKTKYRK  
04eug AIRFDENAAVINKEN-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
05gui SIRFDNAAVINKEN-NPRGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
06mai IIRYDDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
07mar IIRYDDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
08mes IIRYDDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
09nep IIRYDDNAAVIDQKG-NPRGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
10nic IIRYDDNAAVIDQKG-NPRGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
11odo IIRYDDNAAVIDQKG-NPRGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
12oen IIRYDDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
13ory IIRYDDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
14pin MIRSDNAAVIDQKG-NPKGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
15por SIRFDNAAVINKEN-NPRGTRVFGPIARELDRDNFTKISLAQEVVLMSPKTKYRK  
16syn SIRFDNAAVINAEG-NPRGTRVFGPVARELDRDNFTKISLAQEVVLMSPKTKYRK

|       |                                                                |
|-------|----------------------------------------------------------------|
| 01can | HRGRLKGHATKGSALAFGRYGIQALTSSWITSRQIESVRRVIVRHLKRGGKLWIRIFPDK   |
| 02chl | HRGRMRGKASRGNTVVFQDYLQSLQLEASWITSRQIEAARRAMTRQVRRGGKIWIRLFPDK  |
| 03cya | QRGRMGKISTRGNLNVFGDFGLQALEPAWITSRQIEASRRATINRYVRRGGKIWIRIFPDK  |
| 04eug | HRGRLTKIYVLDKVVVFGNYALQSLQLEPWTISRQIEAARRVITRYAKRGGKLWIRIFPDK  |
| 05gui | HRGRLRGIATRGWTLIFGDYGLQALEPIWLTISRQIEATRTTITRQVKRVGLWIRVFPDK   |
| 06mai | HRGRMGKSGCRGNHICFGRYALQVLEPAWITARQIEAGRRAMTRYARRGGKIWIRIFPDK   |
| 07mar | HCGNLKGISTRGNVICFGKFLQALEPSWITSRQIEAGRRATITRYARRGGKLWIRIFPDK   |
| 08mes | HRGKMRGVATRGKNSVFGDFGLKTLEPGWLTISRQIEAGRRAMTRYTRGGQKLWIRVFPDK  |
| 09nep | HRGHRGRALRGNTIAFGDFALQALESSWITSRQIEAARRAMTRYARRGGKIWIRIFPDK    |
| 10nic | HRGRMGKISHRGNHISFGKYALQALEPAWITSRQIEAGRRAMTRNARRGGKIWIRIFPDK   |
| 11odo | HRGRMRGKATRGNEVTFGDYGLQALEPTWITSRQIEAARRTITRYTKRGAALWIRIFPDK   |
| 12oen | HRGRMRGISYRGNRCIFGKYALQALEPAWITSRQIEAGRRAMTRNVRGGKTWIRIFPDK    |
| 13ory | HRGRMGKGSYRGNCICFGRYALQALEPTWITARQIEAGRRAMTRYARRGGKIWIRIFPDK   |
| 14pin | HRGRMGKGSYRGNRCIFGFRALQALEPAWITSQIEAGRRITINRYARRGGKIWIRIFPDK   |
| 15por | HRGRMGKGSASKGNTIAFGDYALQATEPVWLTISRQIEATRTITRYVRRGGKLWIRVFPDK  |
| 16syn | QRGRMRGLAERGSTLNFQDYLQATEPCWITSRQIEAARRAMTRYIRRGKKIWIRIFPDK    |
|       |                                                                |
| 01can | AVTAKPLETRMGSGKSGPEHWIAVVKSGHILFEIDGVSLELAKEAVKLAIKPIKCKFL     |
| 02chl | PVTMRPAETRMGSGKGAPEFWAVVVRPGKVLVELKGVPEVARVALRLAASKLPVKTL      |
| 03cya | PVTMRPAETRMGSGKGAPEFWAVVVRPGKVLVELKGVPEVARVALRLAASKLPVKTL      |
| 04eug | PVTFRAAETRMGSGKGNVEYVVAIVKPGKILYEVLGISEIAKYSKLIAGYKMPIKTRVI    |
| 05gui | SISAKPPETRMGAGKGAPEYVVAIVKPGHILFEINGVSQDLRLYAFKNASYKLPIKTKFI   |
| 06mai | PVTIRPTETRMGSGKSGPEYVVAIVKPGRILEYMGVSETVARAAISIAASKMPIRSQFL    |
| 07mar | PITIRPAETRMGSGKSGPEYVVAIVKPGKILYELISGVSENIARAAMKIAAYKMPIRTQFI  |
| 08mes | PVTIRPAETRMGSGKSGPEYVVAIVKPGTILYEMKGVPTIARASMRVAGFKMPVKTFQV    |
| 09nep | AITMRPAETRMGSGKSGPEYVVAIVKPGKILYEMKGVPTIARASMRVAGFKMPVKTRIL    |
| 10nic | PVTLRPAETRMGSGKSGPEYVVAIVKPGRILEYMGVGTENIARRAISIAASKMPIRTQFI   |
| 11odo | TVTARAESRMGSGKGAVDYVVAIVKPGTILFEIASVPEEIAATALHLASYKLPIKTKFI    |
| 12oen | PVTLRPAETRMGSGKGNPEYVVAIVKPGRILEYMGVGAENIARAKISIAASKMPIRTQFI   |
| 13ory | PVTIRPTETRMGSGKSGPEYVVAIVKPGRILEYMGVSETVARRAISIAASKMPIRSQFL    |
| 14pin | PITMRPAETRMGSGKSGPEYVVAIVKPGRILEYMGVSETVARAAARIAAYKMPIRTQFI    |
| 15por | PVTARPAETRMGSGKGAPEYVVAIVKPGHILFEITGVPKTAQAMKALASYKLPIKTKFI    |
| 16syn | PVTMRPAETRMGSGKSGPEYVVAIVKPGRVFELAGVSEEVAREAMRLAAQLPIKTKFI     |
|       |                                                                |
| 01can | SNDEFAKD---MVRVKRGNVARKRRKKILKLASGFKGAHSHKLFVANQQVKNLSRYSYVG   |
| 02chl | VK-----MTRVKRGNVARKRRKQVLNLASGFRGSSSRLFRTAQQQTKMALSYSYRD       |
| 03cya | SSRV-----MTRVKRGNVARKRRKKILKLASGFRGAHSHKLFVANQQVMKALRYAYND     |
| 04eug | VKI-----MTRIKNNGISKKRRKKISKMGWVGSHSKLFRTCNQQLMKARHYAFYD        |
| 05gui | SR-----MVRIKRGNIAARKRHKKILKLAKGFRGSHSKLFRIANQQVMKALRYGYHG      |
| 06mai | RLEI-----MTRVPRGYIARRRRTKMSFASNFRGAHLRLNRVITQVRRAFVSSHLD       |
| 07mar | TTSSLNKKQEIMTRVKRGYVARKRRKNILTLTSGFQGTSHSKLFRANQQGMRALASSHRD   |
| 08mes | VRDV-----MTRVKRGNVARKHRNKILNLAKGFRGAHSHKLFRTANQIISKLRYAYRD     |
| 09nep | QR-----MTRVKRGYVARKRRNKILANRSFRGTSHSKLFRIANQQHMKALRYSYRD       |
| 10nic | IS-----MTRIKRGYIARRRRTKIRLFASSFRGAHSHKLTTRITQKIRALVSAHRD       |
| 11odo | TRIKI-----MARVKRGNIAARKRAKKILQLAKGYRGAHSHKLFRIANQQVMKALRYSYVG  |
| 12oen | ISG-----MTRIRRGYIARRRRTKTRFFASSWRGAGHNLTRAIIQIRIRAFVSSHLD      |
| 13ory | RLEI-----MTRVPRGYIARRRRAKMSFASNFRGAHLRLNRMITQVRRAFVSSHLD       |
| 14pin | TT-----MTRVKRGYIARKRRKKILAFVSGSRGAHSHKLFRTANQRKARALVSAHRD      |
| 15por | VRNTIES---MSRVKRGNVAKRRRAKFLKAGFKGAHSHKLFRTAQKQVLKALRYSYVG     |
| 16syn | SRQEDYI---MTRVKRGNVARKRRKKILKLAKGFRGSHSKLFRANQQVMKALRNAYRD     |
|       |                                                                |
| 01can | RKLKKRRFRRLWILRLN--AASREEGLN--YSKLVNSLKLRLIQLNRKLSQLAMIDNDAF   |
| 02chl | RQKKKREFCGLWVTRLN--AARLYGLN--YTNFRHNLKAGIQLNRKVLSQLAVRDKQAF    |
| 03cya | RNRKRDFRALWIARIN--ASARLEGMT--YSKLMGSLKKLNIILNRKLSQLAIYDKDAF    |
| 04eug | RKKKNLNLKTLWITRINGLKYLYLTINEKYNIFVSFLRKTITYVNNKLLSEINVRDSKSF   |
| 05gui | RKKRKRREFSLWITRIN--AAVRIEGTN--YSCFINSLKRQHIALNRKMLAQLAVSDQNAF  |
| 06mai | RGRQKRDFFRLWITRIN--AATRVYVNFVNSYSKLIHNLKSKELILNRKMLAQLAVSNPNL  |
| 07mar | RGRKRNLRRLWITRVN--AAARDNGIS--YNKLIYLYKKKILLNRKILAQIAILDKFCF    |
| 08mes | RARKKRDFFRLWIARIN--AASRQNNIS--YSQLINQLKTSNILLNRKILAQIALLDGPVF  |
| 09nep | RACKKRDFFRLWITRIN--AVVRSYGLN--YSRFMHQLRLGNMVLNRKVLSQLASLDPASF  |
| 10nic | RDRKKRDFFRLWITRIN--AVIRERGVSYSYSLIHDLYKRQLLLNRKILAQIAISNRNCL   |
| 11odo | RKQKKRVFRKLWITRIN--AASRLNGLS--YSRLIHNFKSNIELNRKMLSQIAVLDIPTF   |
| 12oen | RTRQKRDFFRLWITRIN--AAIRENGRSSIYSKLIHNLKYLKQLFLNRKMLAQLAILNRNCL |
| 13ory | VRQKRDFFRLWISRIN--AATRIHKVFDNYSKLIHNLKYLKELILNRKILAQVAVLNSNNL  |
| 14pin | RGRKRDRLRLWITRIN--AAARANGVS--YNNFIQYLYKRQLFLNRKTLAQIAVLDSNCF   |
| 15por | RKKKKRDFFRLWITRIN--AAHSQGMN--YSTFITALKDNIALNRKMLAQLASNDIKAF    |
| 16syn | RKKKKRDFFRLWITRIN--AAARQEGMS--YSKLTGQLKKANIEINRKMLAQLAVLDPAAF  |
|       |                                                                |
| 01can | KRLVASSRV-----MAIKKYKPYTPSMRGRVLASDFFDLSEKKAPKRLSFGI           |
| 02chl | EQLILLVKD-----MAIRFFKAATPGTRHGSVLD--FSEITHKKPEKALTSWW          |
| 03cya | MEILKTIP-----MAIRSYKATPGTRNRTISE--FSEITKSEPEKSLTLFK            |
| 04eug | SHLSKPIKSTGINL-----MIIRYKPYTSGTRKRSVSN--FSEITKSNPEKYLTFFV      |
| 05gui | KQLTKITH-----MGIRIYKSYTPGTRNRSSTD--FVEITKSKPEKSLLRKK           |
| 06mai | YTTISNK--IRIIN-----MAKHLKYTPIPSTRKGTV----DRQVKSNNRKNLIHGR      |
| 07mar | STIIKNIITE-----MAIRLYRAYTPGTRNRSVVP--FDEIVKQCPQKKLTYNK         |
| 08mes | SQIVMESNS-----MGIRLYKATPGTRNRSVLE--FNDITKTNPESKSLTYHR          |
| 09nep | NRLIRAT-----MGIRFYRAHTPGTRNRSVSD--FHEITSTPTKSLTHAN             |
| 10nic | YMSINEIIEVDWKESTRII--MAIHLKYTSTPSTRNGTV----DSQVKSNNRNNLIYQG    |
| 11odo | NQLIAISK-----MSIRLYKSYTPGTRNRALSS--FTEITKTKPEKSLIQKN           |
| 12oen | YMSINQILKEVDWQESATILEIMAMHLKYTSTPSTRNGAV----DSQVKSNNRKNLIYQG   |
| 13ory | YTTISNK--IKIIN-----MAKHLKYTPIPSTRKGTI----DRQVKSNNRKNLIHGR      |
| 14pin | STIFNNLSYDEIG-----MAIRSYRLTPTDRNRSVSG--FDGRVQLDPQKKLTSSGQ      |
| 15por | QAILETVKN-----MAIRLYRAYTPGTRNRTVST--FSEITTDKPEKSLINKH          |
| 16syn | SEVVVKAATAK-----MGIRNRYMPTGTRQASVSD--FTEITKSKPEKSLTTNR         |

```

01can K-SISGRNNQGGKITCRHGGGGHKRYRLVDFKRCCTGVLAQVSDIYDPNRSAHIALNY
02chl S-RSKGRNNRGIITSRHRGGGKHRLYRKIDFARAKVNPVAKVAYIEYDPNRRNARIALVNY
03cya H-RKKGRRNNGIITTAHKGGGGSKRLYRKIDFRRDLKLVPAKVAAIEYDPNRRNARIALHY
04eug H-RSKGRNRSRGVITCRTLGGGKHRLYRIEFKRNKLGILKVISIEYDPNRRNARIALIYY
05gui L-SCAGRNNRGLITVRHGGGGHKRYRLVDFKRNKLDIPAIVASVEYDPNRRNARIALHY
06mai HRCGKGRNARGIITARHGGGKHRLYRKIDFRRNQKDISGRIITIEYDPNRRNAYICLIHY
07mar H-IKKGRRNNGIITSQHRGGGKHRLYRKIDFRRNKYITGKIKTIEYDPNRRNAYICLINY
08mes H-RSKGRNNRGIITIRHGGGKHRLYRKIDFRTREK-NIPATVASIEYDPNRRNARIALYY
09nep H-RARGRNHSGSITTRWRGGGKHRLYRKIDFRRDKVGVLARVATVEYDPNRRSARIALHY
10nic HHCGKGRNARGIITARHGGGKHRLYRKIDFRRNEKDIYGRIVTIEYDPNRRNAYICLIHY
11odo H-RSKGRNNRGVITIRHGGGKHRYRIIDFGRKKHNVEGVAAIEYDPNRRNARIALHY
12oen PRCSKGRNARGIITAGHRGGGKHRLYRKIDFRRNERDIYGRIVSIEYDPNRRNASICLIHY
13ory HRCGKGRNRSRGITARHGGGKHRLYRKIDFRRNQKDISGRIIVTIEYDPNRRNAYICLIHY
14pin HHCGKGRNNGRGITARHGGGKHRLYRKIDFRRNKEHISGEIVTIEYDPNRSAYICKVHY
15por H-FCKGRNNRGVITCRHKGGGHKRYRLIDFRRNRHNIIAKVASIEYDPNRRNARIALHY
16syn H-DQKGRNNRGVITSRHRGGGKHRLYRKIDFRRNKQNPARVAAIEYDPNRRNARIALFY

01can LDGEKSYIISPMLKVGTYVVSQKEASPDIGNALPNCVPLGFEIHNIELIHGKGGQVAR
02chl QDGEKKYILHPVGLVQVQTIIASPEASIAIGNCLPLVKIPLGTEVHNIELQPSGGQLVR
03cya QNGEGYILHARGLA VGNMVSQGNAPIEVGNLPLSEIPLATEIHNIELTPGKGGQLVR
04eug KNGDKSYIHPFDLCVGNIIISDFFSPIKIGNSLPISKIPLGTIIHNVEFEPGKGGQIAR
05gui QDGEKRYILHPKLA VGDKIYSGINVP IEGNAMP LYNVPLGTAVHNVELIPCRGGQIVR
06mai GDGEKRYILHPRGAIIGDTIVSGTKVPI SMGNALPLTDMPLGTAIHNIEITRGGGQLAR
07mar EDGEKRYILYPRGILKDDTIISSEAPILIGNTLPLTNMPLGTAIHNIEITPGKGGQLVR
08mes KNGIKRYIHPRGLSVGKEIVSSVEAPLSVGNLPLNKIPLGTGIHNIELSPGKGGQLAR
09nep QDGSKRYILHPQGLAIGAEVMSSEAPISIGNALPLVNMPLGTEVHNIELRPYNGGQLVR
10nic GDGEKRYILHPRGAIIGDTIVSGTEVP IKGNGALPLTDMPLGTAIHNIEITLKGKGGQLAR
11odo TDGEKRYILHPNNLVGDRVVSQMEAEI VIGNALPLEKIPLGASVHNIELIPNRGGQIVR
12oen GDGEKRYILHPRGAIIGDTIVSGTEVP IKGNGALPLTEMPLGTAHNIENTLKGKGGQLAR
13ory GDGEKRYILHPRGAIIGDTIVSGTKVPI SMGNALPLTDMPLGTAIHNIEITRGGGQLAR
14pin KNGDKMYILHPRGVMIGDTILSGPKAPISIGNALPLTNMPLGTAIHNIEITLKGKGGQLAR
15por LDGEKRYILHPRSLVGAIIVSSGPMAPIEVGNALPLSTIPLGTAVHNIELRPYCGGQIVR
16syn TDGEKRYILAPAGLVGMTVIAGEEAPFEIGNTLPLSRIPLGSEIHNVELVAGRGQMVVR

01can AAGTSAKLIASQDYVTIKLPSGEIRLFRGECYATIGKVGNIHDHNEKIGKAGNRWLGI
02chl AAGTVAQIVAKEGTWASRLRPSGEVRLVSQNCWATIGRVGNIDAFNLTLGKAGRSRWLGR
03cya SAGSSAQLLAKEGNYVTLRLPSGEMFRVRECYATIGQIGNAEISNISIGKAGNRWLGI
04eug AAGTFVQILANEGKFVTITMPSGEVRLRLRYCWATIGQVGNLDHNSVNLGKAGNRWLGN
05gui SAGTSAQVVAQDGGVVTIKMPSNEVMIYKNCYATIGEVGNADIKNIRLGKAGRSRWLGI
06mai AAGAVAKLIAKEGLATLRLPSGEVRLVSQNCLATVGQVGNVGNQKSLGRAGSKCWLGR
07mar AAGTVAKIIAKEGQVTLRLPSGEIRLISQKCLATIGQIGNVDVNNLRIGKAGSKRWLGK
08mes AAGAVAQLIAKEGKFVTLRLPSGEVRLILKECWATIGQVGNVDANNITIGKAGTRWLGR
09nep AAGAVAQLVAKEGGFGLRMPSGEVRLVAKDCWATVGQVGHVESINLTLGKAGRSRWLDR
10nic AAGAVAKLIAKEGKSATLRLPSGEVRLISKNCSATVGQVGNVGNQKSLGRAGSKRWLGK
11odo AAGTSAKILAKEGDYVTLRLPSKEIRLIRKECFATIGEVSNNDAPLVQSGKAGRSRWLGR
12oen AAGAVAKLIAKEGKSATLRLPSGEVRLISNNCSATVGQVGNVGNQKSLGRAGSKRWLGK
13ory AAGAVAKLIAKEGKSATLRLPSGEVRLVSQNCLATVGQVGNVGNQKSLGRAGSKCWLGR
14pin AAGAVAELIAKEDRSATLRLPSGEVRLISENCSATIGQVGNITANNRSFGKAGAKRWLGK
15por SAGTYAQIVAKEGNFVTIKLPSSEVRMIRKECYATIGQVGNIDASNITLKGAGRSRWLGR
16syn SAGAFQAQVAKEGDYVTIKLPSKEVRMVRKECVATLGRVSNAEFRNLKLGKAGRSRWLGR

01can RPTVRGSAMNAVDPHPHGGG-EGRSPIGRSQPSTPWGRPALGIKTRRNK-FSNFYILRRRK
02chl RPHVRGSAMNPVDHPHGGG-EGRAPIGRARPVSPWGRPALGAKTRKRKKFSAALILQRRK
03cya RPTVRGVVKNPVDHPHGGG-EGRAPIGRSTPVTWPWKPALGRTRARTTKYSDNLIIRRRK
04eug KPTVRGVAMNPVDHPHGGG-EGRSPIGRPKPVTWPWKPALGKKTRSPKRFSNKYIIRSRK
05gui RPSVRGVVKNPVDHPHGGG-EGRSPIGRAPKPTWPWKPALGVGTRQNKYSDFCIIIRSRN
06mai RPTVRGVVKNPVDHPHGGG-EGKAPIGRKKPTTPWGPALGRTRTRKKKYSDFILRRRK
07mar RPKVRGVVKNPIDHPHGGG-EGRAPIGRKKPLTPWGHGALGKRSRKNKYSDTLILRRRK
08mes RPTVRGVVKNPVDHPHGGG-EGRSPIGRPKPVPSPWGTALGAKTRKRKKYSVLIIRRRK
09nep RPRVRGVVMNACDHPHGGG-EGRCPIGHPGLTPWGPALGQRTARKKYSALLVRRRK
10nic RPTVRGVVKNPVDHPHGGG-EGRAPIGRKKPTTPWGPALGRRSRKNKYSDNLIIRRRS
11odo RPTVRGVVKNPVDHPHGGG-EGRAPIGRTRPLTPWGPALGIKTRKNKASDAYILRRRS
12oen RPTVRGVVMTPTVDHPHGGG-EGRPPIGRKKPATPWGSPALGRRSRKNKYSENLIIRRRS
13ory RPTVRGVVKNPVDHPHGGG-EGKAPIGRKKPTTPWGPALGRRTRKRKKYSDFILRRRK
14pin RSEVRGVAMNPVDHPHGGG-EGRTPIGRKKPTTPWGSALGKRSRKNKYSDASILRRRE
15por RPTVRGVVKNPVDHPHGGGEGKSPIGRSRPVTWPWKPALGVKTRNPNKYSNPYVLLVNN
16syn RPHVRGVVKNPVDHPHGGG-EGRAPIGRSGVPSPWGPALGAKTRKNKRSALIVRRRT

01can -----MKVRASVRKICSR---CVALKRHGVLIIVLCS-NPKHKQRQG-
02chl -----MKVRPSVKKMCCK---CRLIKRKGTLRIVICQ-NPKHKQRQG-
03cya -----MKVRASVRKMCEK---CRTIRRGVRMVIC-SNPKHKQRQG-
04eug MV-----MKIRSSVKKICNK---CYLIRRNKNNLIVCI-NNPKHKQRQG-
05gui -----MKVSSSIGSLKNRSKDCQIVKRRGRIVYICKTDPLRKVRQGG-
06mai -----MKIRASVRKICTK---CRLIRRRGRIRVICS-NPKHKQRQG-
07mar NS-----MKIRASVRKICEN---CRLIRRRRRIMVICS-NPKHKQRQG-
08mes SA-----MKVRASVRKICEN---CRLIKRRGTVMIVICSNNPKHKQRQG-
09nep -----MKVRPSVRKICDK---CCLIRRRHKLIVICS-NPKHKQRQG-
10nic K-----MKIRASVRKICEK---CRLIRRRGRIVICS-NPRHKQRQG-
11odo -----MKVRPSVKKMCCK---CRVIRKRGKIMVICP-NAKHKQRQG-
12oen K-----MKIRASVRKICTK---CRLIRRRGRIVICS-NPRHKQRQG-
13ory -----MKIRASVRKICTK---CRLIRRRGRIRVICS-NPKHKQRQG-
14pin -----MKIRASIRRICGK---CRPIRRRKRVMIICS-NPRHKQRQG-
15por KVLTYNLIKLYNEINTMKVRPSVRKMCEK---CRIIRRRHGRMVICN-NPKHKQRQG-
16syn K-----MKVRASVKKMCCK---CRVIRRRGRMVICSNANPKHKQRQG-

```

```

01can -----MLQSKKSSKK-----LKKSVIDGLIIINSTFNNTIVTATDC
02chl -----MAKKIEKSAKK-----FREKVLGVVAHIQSTFNNTIVTITN
03cya -----MARQIKRSGTTK-----QKKNIIPGVVAHIQSTFNNTIISITS
04eug -----MIKQKKKI I-----TKKKFPKGI VHIKTTYNNVLVSVD
05gui AKMKRKMVRQVKKSSLLK-----NKKVAANGVVHIQSTFNNTIVTSTL
06mai -----MTKAIPKIGSRKKVRIGLRRNARFSLRKSARRITKGIHVQASFNNTIITVTD
07mar -----MPKSVKKINLR-----KGKRRLPKGVVHIQASFNNTIVTVDI
08mes -----MAKQIRKIGVR-----KTKRKIPKGVVHVQATFNNTIVTIDV
09nep -----MARQSKRLAPRR-----LKRKVSQGVVHIQASFNNTIVTIDS
10nic -----MAKAIPKISSR-----RNGRIGSRKGARRIPKGVVHIQASFNNTIVTVDV
11odo -----MAKQIRKSTFKK-----DNSSINGVVHIQSTFNNTIVTITNI
12oen -----MAKSIPASGLR--LRLRLRNARRRSRSTRKIPKGVVHIQASFNNTIVTVDV
13ory -----MTKAIPKIGSRKKVRIGLRRNARFSLRKSARRITKGIHVQASFNNTIITVTD
14pin -----MSKTIKRIGSR-----RNEHRVLKGVIVVQASFNNTIVTATDV
15por -----MARQIKKSGARK-----NKHNAVNGI THIKSTFNNTIVTITNL
16syn -----MARPTRKTGPKK-----AKKNVPSGVAHIQSTFNNTIVTISDI

01can YGKVIASAGSGTEGFGAKKGTFFAAQSATEKL I KALLEQGTQRIHISVSGPGPRETSI
02chl KGNVLAWSSAGACGFGKARKTPLAAKQAAENAAQTCVSQGMREIRVNVKGAGAGREAAL
03cya TGEVIAWASAGSSGFGKARKGTPFAAQTAAENSARQAMEQGMQRIEVIISGPGSGREMAI
04eug KGNVIAWSSAGACGFGKSKKATPLATKTTTIAVKKAEIQLKQVEINISGPGTGRETAL
05gui EGDVTAWASSGAIGFGAKKGTFFAAQIAAEKATKEAISQGMKKAIEVLINPGSGRETAI
06mai QGRVVFWSAGTCGFKSSRSKASPYAGQRTAVDAIR---TVGLQRAEVMVKAGSGRDAAL
07mar RGQVVSWSAGACGFGTKKSTPFAAQTAAENAIRLIDQGMKQAEVMIISGPGGRDATAL
08mes RGEVLWSSAGACGFGTKKGTFFAAQTAAENAVRQVIDQGMQAEIMISGPGSGRETAI
09nep SGGVLAWSSAGACGFRGAKKGTFFAAQTAAETAIRQCIDQGMQADIIIVQPGNGRETAI
10nic RGRVVSWSAGTSGFGKTRRGTPFAAQTAAANAIRVVDQGMQRAEVMIKGPGGLGRDAAL
11odo TGDTSWASAGSSGFGKARKGTPFAAQTAAEKAALDALSKGMKSVEVLVKGQSGRETAI
12oen RGRVVSWSAGTCGFKSTRKGTFFAAQTAAAGDAIRPVVDQGMQRAEVIKGPGLGRDAAL
13ory QGRVVFWSAGTCGFKSSRSKASPYAGQRTAVDAIR---TVGLQRAEVMVKAGSGRDAAL
14pin RGQVLSWSSAGACGFGTKRGTFFAAQTAAENVIRALMDRGMERVEVMISGPGGRDATAL
15por KGETLSWSSGASGFGKAKKGTFFAAQTAAEKAQAMDQGMQRETVLVNPGPGAGRETAI
16syn RGDVISWASAGSSGFGKAKKGTFFAAQTAAASAARRAMEQGMRLQEVMSVSGPGAGRETAI

01can RSFQTSGLQIISIKDITSVPFNGCRPPKRRIMPTIQLIRFERKQIVGTSKSAALESCP
02chl RGLRDAGLNTTIIIRDITP IPHNGCRPPKRRIMPTIQLVRSARTL SKKTKSPALKSCP
03cya KALQATGLEISLIRDITPVPHNGCRPPKRRVMPTIQLIRSKRTKIEKTKSPALKACP
04eug KCVQSLGLRISCIRDVTP LPHNGCRPSKRRIMPTLEHLTRSPRKKIKRKTSPALKGCP
05gui RALQAAGLEITLIRDITPVPHNGCRPPKRLRVMTIQLIRSERQTI--QKTKSPALKSCP
06mai RATAKSGVRLSCIRDVTPMPHNGCRPPKRRIMPTVKQLIRNARQPIRNARKSAALKGCP
07mar RAIIRSGIILSFVRDVTMPHNGCRPPKRRVMPTIQLIRNKRPPIENRTKSPALKGCP
08mes RAIQAAGLGITLIRDVTP IPHNGCRPPKRRVMPTIQLIRSERKRVNNTKSPALKACP
09nep RALQLAGVGSILIRDITSVPHNGCRPPKRRIMPTIQLVRAQKRRI VKTKSPALCACP
10nic RAIIRSGIILLTFVRDVTMPHNGCRPPKRRVMPTIKQLIRNTRQPIRNVTSPALRGCP
11odo RSIEGAGFEVTSIQDITPVPHNGCRPPKRRVMPTIQLVRSRRIQIKKTKSPALVNCP
12oen RAIIRSGIRLSCIRDVTP LPHNGCMPKPKRRV-----
13ory RATAKSGVRLSCIRDVTPMPHNGCRPPKRRIMPTVKQLIRNARQPIRNARKSAALKGCP
14pin RTIRRSIGILLSFVRDVTMPHNGCRPPKRRVMPTIQLIRNARQPIENRKKSPALRGCP
15por RALQAAGLEITLIRDITPVPHNGCRPPKRRVMPTIQLVRSERRIKKTKSPALKSCP
16syn RALQAGGLEITLIRDVTP IPHNGCRPPKRRVMPTIQLIRSERKSVKQKTKSPALKGCP

01can QKKGVCVKVYTTTPKKPNALRKVARVRLTSGFEITAYIPGIGHNLQHSVVLIRGGRVK
02chl QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGIGHNLQHSVVLVRGGRVK
03cya QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGIGHNLQHSVVLVRGGRVK
04eug QKRAICMRVYTTTPKKPNALRKVARVRLSSGLEVTAIPGIGHNLQHSVVLIRGGRVK
05gui QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGIGHNLQHSVVLIRGGRVK
06mai QRRGTCAVYTTIPKKPNALRKVARVRLTSGFEITAYIPGIGHNLQHSVVLVRGGRVK
07mar QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEITAYIPGIGHNLQHSVVLVRGGRVK
08mes QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGIGHNLQHSVVLIRGGRVK
09nep QRRGVCTRVYTTTPKKPNALRKVARVRLSSGFEVTAIPGIGHNLQHSVVLIRGGRVK
10nic QRRGTCTRVYTTIPKKPNALRKVARVRLTSGFEITAYIPGIGHNLQHSVVLVRGGRVK
11odo QRRGVCTRVYTTTPKKPNALRKVARVRLTSDLKVTAYIPGIGHNLQHSVVLIRGGRVK
12oen -----MGGVRFKIVRPAPTVPVYAST-GITHKGLRQ--T--SGKMPA
13ory QRRGTCAVYTTIPKKPNALRKVARVRLTSGFEITAYIPGIGHNLQHSVVLVRGGRVK
14pin QRRGVCAVYTTIPKKPNALRKVARVRLTSGFEITAYIPGIDHNLQHSVVLVRGGRVK
15por QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGVGHNIQHSVVLIRGGRVK
16syn QRRGVCTRVYTTTPKKPNALRKVARVRLTSGFEVTAIPGIGHNLQHSVVLIRGGRVK

01can DLPGVRYHIIRGALDSTGVKNRLRARSKYGASKPKK--MAKKSIVIQRNINRLKLINKYSA
02chl DLPGVRYHIIRGTLDSAGVKDRSQRSKYGVKRPK--MAKKSMIERDRKARLITKYAA
03cya DLPGVRYHIIRGALDAAGVKDRQSRKYGAKRPKA--MAKKSMIERDKRENLMNKYLV
04eug DLPGVKYHVIRGCLDAASVKNRKNARSKYGVKKPKPK-MSKKSIIARQKRIILVLIHSH
05gui DLPGVRYHIIRGTLDAAGVKNRQSRKYGAKPKPE--MAKKSMIERERKREELVSKYEK
06mai DLPGVRYRIIRGTLDAVAVKNRQQRSKYGAKKPKK--MAKKSIIQKQKQKLEQKYHL
07mar DLPGVRYHIIRGTLDAVGKDRQQRSKYGVKKSK--MAKKSIIQREKQKQLEKYYKI
08mes DLPGVRYHIIRGILDTAGVKDRQQRSKYGVKRS--MAKKSMIERKQKQKLVNKYAV
09nep DLPGVRYHIIRGTLDTAGVKERQQRSKYGVKRVTKK-MAKKSMIERESKRAALVAKYAT
10nic DLPGVRYHIIRGTLDAVGKDRQQRSKYGVKKPK--MARKSLIQREKKRQKLEQKYHS
11odo DLPGVRYHIIRGALDSVGKDRQQRSKYGVKKPKSDKMAKKSMIERKKRIKLNKYYTP
12oen RNPADKLQYIVQGLSTYIRDFGTGRSQNGGTIG---MARKGLIQREKKREKLEQKYRL
13ory DLPGVRYRIIRGALDAVAVKNRQQRSKYGVKKPKK--MAKKSIIQREKKRQKLEQKYHL
14pin DLPGVRYHIIRGTLDAAEVKDRQQRSKYGVKKQK--MARKSLIQREKKQALERKYHL
15por DLPGVRYHVIRGTLDAAGVKDRKRSRKYGTKKPKS--MAKKNMIQREIKREKLEKYYL
16syn DLPGVRYHIIRGTLDTAGVKDRQQRSKYGTREKAKKMAKKSMIERDKRRSLVAKYAA

```

|       |                                                               |
|-------|---------------------------------------------------------------|
| 01can | QREAIKNEIK---RTSKVEKKISLYSNIISRLPRDSSKVRLRSRCWVTGRGRSVYKNFGLS |
| 02chl | KRNLLVEIK---TATSLEDKFNLHRKLQQLPRNSAPVRSNHRCTITGRPGYFRDFGLS    |
| 03cya | KRQLKTRLN---ETDSVEEKLLNLQELQKLPNSAPSRVIRCWLTGRPRGNYRDFGVS     |
| 04eug | NRVYVRTNGK---DEKSFEKKLRIYSFLQKLPNLSRLCRNLNRCYVTGRSGYFRFTGLS   |
| 05gui | KRLKLSKLG---KTAEEYEKLEVYKKIQEIPRNAPPRLNRCWVTGRSGYFRDFGLS      |
| 06mai | IRRSSKKKIRSKVSPLSLSEKTKMQEKLQSLPRNSAPTRLHRCFLTGRPRANYRDFGLS   |
| 07mar | LRNSLKKKIT---ETSSLDEKWEFKKLQSLPRNSAPTRLHRCFLTGRPKANYRDFGLS    |
| 08mes | KRKELKEQIK---TSVSFEERFKLQELQKLPNNSPTRLHNRCSVTGRPKGYRDFGLS     |
| 09nep | RRQALKAQI---KTKSFDERLVLQHQLQDLPVNSVPCRLHNRCTITGRPKGYRDFGLS    |
| 10nic | IRRSSKKEIS---KVP-SLSDKWEIYKGLQSPPRNSAPTRLHRCFLTGRPRANYRDFGLS  |
| 11odo | KRNTLLQAYR---QTEDFQSRDLIHSKIQKLPNRSAPNRIRNRCWVTGRPGYFRDFGVS   |
| 12oen | IRRSSKKEIS---TDP-SLSEKWKIHGELQSSPRNSATMTSSRCFSTGRPVNYRDFGLS   |
| 13ory | IRRSSKKKIRSKVYPLSLSEKTKMREKLQSLPRNSAPTRLHRCFLTGRPRANYRDFGLS   |
| 14pin | IRQSLEEKSK---VSSLDKWEIHRKLQSSPRNSAPTRLHRCFSSTGRPRANYRDFGLS    |
| 15por | KRLAIKEQLK---KTTSAEKIELRQKLQEMPRNSAPVRSRNRWLTGRSGYFRDFGLS     |
| 16syn | KREALKEEFR---QAETLEDKLAVHQQLDLPNRSAPNRRNRNRCQVTGRPSYRDFGLC    |
|       |                                                               |
| 01can | RHMFRFMASNGLLPGVVKSSW-----MNKKI                               |
| 02chl | RHVLREYALQGFPGVVKASW-----MAGIQKRKTIKSFKTRRK                   |
| 03cya | RHVLREMAHQGLLPGVTKSSW-----MSVYRRR                             |
| 04eug | RHILDRMAHYGLLPGVTKASW-----M                                   |
| 05gui | RHVLREMVHDCLLPGVTKSSW-----MAIFRRK                             |
| 06mai | GHILREMYACLLPGATRSSWYISKQPFRRSKQPFRRSKQTFHKSQKQPFRRKQPFRRK    |
| 07mar | RHLLREMAHACLLPGVTKSSW-----MNKSK-----RSSRRR                    |
| 08mes | RHVLREMAHECCLPGVTKSSW-----MTFSRRR                             |
| 09nep | RHELREMAHGCCLPGVTKASW-----MIKFSRRR                            |
| 10nic | GHILREMVHACLLPGATRSSW-----MDKSKRPFKFKRSFRRR                   |
| 11odo | RHVLREMAHSCCLPGVTKSSW-----MLAQKQK                             |
| 12oen | GHILREMVQACLLPGATRSSW-----MDKSKRFLKSKRSFRRR                   |
| 13ory | GHILREMYACLLPGATRSSWYTSKQPFHKSQTFHKSQKQTFRRKQKQPFRRKQPFRRK    |
| 14pin | GHILREMAHACLLPGIKKSSW---MKQTMKPKQSFRRHFKPIRRRFKPIRRYLPKPIRRH  |
| 15por | RHVFREMSHECCLPGVTKSSW-----MAVYRKK                             |
| 16syn | RNVLEWAHQGLLPGVTKSSW-----MNYRKR                               |
|       |                                                               |
| 01can | KIPKD-----SKAINYKTTDLIREFISDRGKIIPQRTKS--TRKKHK               |
| 02chl | VVPVLIPKKGQAVISNTGEPASRYIIDYKNTQLLVKFISPGQKILSRRTATGL--TAKQQR |
| 03cya | LSPLK-----P-----NQIDYQDVLLRFTITDQKILPRRTVGL--TAKQQR           |
| 04eug | KYP-----ALIDYKNVILRRFINFQKIIIPKRLNKPCLTYKQHR                  |
| 05gui | NSPIK-----K-----DDFIDYKDVLLSKFLTEQKILPRRTITGL--TMKQQS         |
| 06mai | KQPFRRKSKQPFRRRSRI---GPGDRIDYRNMSLINRFISEQKILSRINRL--TLKQQR   |
| 07mar | MPPIR-----S-----CEIIDYKNISLLRRFVSEQKILSRMNRL--TSKQQR          |
| 08mes | TSPIK-----P-----TDSIDYKNIDLLSQFITEQKILPRRVNRI--SAKQQR         |
| 09nep | LAKLAN-----AQ-----IDDRIDYKNVDLLRFLTEEGKILPRRTVGL--TAKQQR      |
| 10nic | LPPIQ-----S-----GDRIDYRNMSLISRFISEQKILSRVNRL--TLKQQR          |
| 11odo | LSPIK-----V-----NQIDYKDIDLLKFLITEQKILPRRTATGL--TVQQQR         |
| 12oen | LPPIQ-----S-----GDRIDYRNISLISRFISQKILSRVNRL--TLKQQR           |
| 13ory | KQPFRR-----RRPRI---GPGDRIDYRNMSLINRFISEQKILSRINRL--TLKQQR     |
| 14pin | LSPIR-----S-----GDRIDYKNMSLISRFISEQKILSGVRNRL--TSKQQR         |
| 15por | LSPIK-----P-----TEAVDYKDIDLLRKFITEQKILPKRSTGL--TSKQQR         |
| 16syn | LSPLP-----P-----NQPIDYKDTLLRKFITERGKILPRRTITGL--TAKQQR        |
|       |                                                               |
| 01can | LIEKAIKQARLIGLMPFKQKNW-----                                   |
| 02chl | IMANAIKRARMGGLVPFVNYELGSKK-----                               |
| 03cya | AVTKAIKQARVLALLPFVNRES-----                                   |
| 04eug | LLRSVKQARYLGLLPFKTKDFF-----                                   |
| 05gui | RLTKAVKQARILSLLPFINRDMIL---QQAQSIITGPRHRKNRHIPQLTKQFNSNRNLRN  |
| 06mai | LITLAIKQARILSFLPFRNYENEKQFQAQSIITGPRHRKNRHIPQLTKQFNSNRNLRN    |
| 07mar | LLTIAIKQARVLALLPFLNNEN-----                                   |
| 08mes | AITKAIKQARLTLPLFLNQEI-----                                    |
| 09nep | GMTRAIKQARIMALLEFIHR-----                                     |
| 10nic | LITLAIKQARILSLLPFLNNEKQFERTESTARTTGFKARNK-----                |
| 11odo | QIAKAIKRARVLSLLPFVASNSI-----                                  |
| 12oen | LITIAINQARILSLLPFRPKAQRFKRSQSTARTVGLRTRNK-----                |
| 13ory | LITLAIKQARILSFLPFRNYENEKQFQAQSIITGPRHRKNRHIPQLTKQFNSNRNLRN    |
| 14pin | LMTNAIKRARILSLLPFLYNNEN-----                                  |
| 15por | KLTKAIKQARILSLLPFLNKD-----                                    |
| 16syn | DLTTAVKRSRLVALLPFVNKEI-----                                   |
|       |                                                               |
| 01can | -----MSRSIKKGPYVHFLLKNTNNLNSNSKRVIKTWSRSSVILPSM               |
| 02chl | -----MARSLKKAPFVANHLLKVERLNTQGDVKTWSRSSSTIVPLM                |
| 03cya | -----MARSLKKGPFIHLLKVELLNTSGKTEVIKTWSRSTILPMM                 |
| 04eug | -----MSRSLKKGPVFFYSILKKVDQMNSNRFKSVIITWSRSTIIPIM              |
| 05gui | -----MSRSLKSGPYIAAHLKKLNNVDIQPDVVIKTWSRSSSTILPMM              |
| 06mai | SNQNLNRNNNRNLSDDCMTR-KKTNPFFVARHLLAKIEKVNKKEKEIIVTWSRASSILPAM |
| 07mar | -----MTRSLKKGPVADHLLKIEFLNKKKEKEIIVTWSRSTIVPTM                |
| 08mes | -----MARSLKKGPVADHLLKIEFLNKKKEKEIIVTWSRSTILPIM                |
| 09nep | -----MPRSLKKGPVFNHLLKIEKMNKGKKEHVIITWSRSTIIPIM                |
| 10nic | -----VTRSLKKNPFFVANHLLKIDKLNKAEKEIIVTWSRSTIIPIM               |
| 11odo | -----MPRSLKKGPVAYHLLKIDKMNASGKDVITVTSRSTILPTM                 |
| 12oen | -----MKRSLKNNPFFVANLLRMEKLNREDKILIRTSRSTIILTM                 |
| 13ory | SNQTLNRNTNRNLSDDCVTR-KKTNPFFVARHLLAKIEKVNKKEKEIIVTWSRASSILPAM |
| 14pin | -----MARSLKKNPFFVANHLLKIEFLNKKKEKEIIVTWSRASSIVPAM             |
| 15por | -----MSRSIHKGPFIIDVSLTRIEALNTSGKKEVIKTWSRSTIIPDM              |
| 16syn | -----MGRSLKKGPVFAASLLRKIDKLNKGDQVVKVITWSRSTILPQM              |



01can VGHTIAVHNGKIHPVIFISDQMVGHKLGEFAFTRSFRRS--HAKIDKKIRK-----MET  
02chl IGHTIAVHNGREHIPVFIITDQMVGHKLGEFAPTRTFRG--HVKKDKKSKR-----MLQKV  
03cya VGHTIAVHNGRQHLPVFIITDQMVGHKLGEFAPTRTFRG--HTKSDKKARR--MA----IV  
04eug IGNTIGVYNGKEHIPVLVSDQMI GHKLGEFVQTRNRYG--HKHDKKTKTKR-----MI  
05gui VGATIAVYNGKQHPVYISDQMVGHKLGEFSPTRTFRS--HIKSDKKAKR--M----TI  
06mai VGHTIAIHNGKEHIPYIITNPMVGRKLGEFVPTRHFTSYESTRKDTKSRR--MTRRYWNI  
07mar IGHTIAVHNGREHLPYIITDQMVGHKLGEFAPTRTFRG--HAKNDKKSRR--MKQKSWNI  
08mes IGHTIAVHNGREHLPYIITDQMVGHKLGEFSPTRTFRG--HTKSDKKARR--MK----SV  
09nep IGHTIAIHNGKSHLPYIFINDRMI GHKLGEFVLTNRNRYG--HGKTDKKARR-----MA  
10nic IGHTIAIHNGKEHLPYIITDSMVGHKLGEFAPLTNFR--CHAKSDNRSSRR--MTRRYWNI  
11odo VGHTIAVYNGRQHPVIFISDQLVGHKLGEFVSTRTFKS--HIKTDKKTKR--MA----DI  
12oen IGHTIAIHNGKEHSPYIITDYMVGHKLGEFAPTNFH--EHAKNDKKSRR--MTRRYWNI  
13ory VGHTIAIHNGKEHIPYIITNPMVGRKLGEFVPTRHFTSYESARKDTKSRR--MTRRYWNI  
14pin IGHTIAVHNGREHLPYIITDQMVGHKLGEFAPTLFQ--HARNDKKSRR--MSRRYWNI  
15por IGHTIAVYNGKQHPVVFVSDQMVGHKLGEFVPTRTFRT--HVKGDRKARR--MA----IV  
16syn VGHTIAVHNGRQHPVVFVSEQMVGHKLGEFAPTRTFRS--HSKSDKKARR--MP----VV

01can SLSELLDASAHIGHSTSKWNPKFFFIHMTKKGIIHIDLIKTIIQISVACMFLKKEVRNG  
02chl KLEDMIQSGMHFGHSTREWNPRMAPYIYGERNGRHILDVQTYLLNKVLAFLQEAQAQ  
03cya TLEDMLEAGVHFGHQSRWNPMSQYIYTERNGIHIIDLVTQAQLLAEAYDFLISAEEG  
04eug TVEKMLNSSVHLGHKVKQWNPMBRIYIYGERKGLHIIDLQTIIVCLLKACNFIIRSVRG  
05gui TLAELLESVGHFGHQAARRWNPMPFYIYAEKNGIHIIDLVTQAQLLAECEVFKSAEAG  
06mai NLKEMIEAGVHFGHGKKNWPKMAPYISAKRKGTHITNLARTARFLSEACDLVFAASQ  
07mar HLEEMMEAGVHFGHQARKWNPMPYIYTERKGIHIINLTQTARFLSEACDLVANASSG  
08mes GLEEMMEAGVHFGHQTRWNPMSYIYGNRNGIHIIDLVTARLLSEACDLFNAKEG  
09nep TLEQMLDVGVHFGHQARKWNPMPYIYGERKGIHIDLQTAILLDKARVYARRQAASN  
10nic NLEEMMEAGVHFGHGTARKWNPMPYISAKRKGITHITNLTRARFLSEACDLVFAASRG  
11odo TLAQLLEAGVHFGHQARKWNPMPFYIYTERNNIHIIDLVSQAQLLEANSYLSAAEKG  
12oen NLEEMMEAGVYFGHGKKNWNPMPYIYANRKGITHITNLTKARFLAKACDLVFAASRG  
13ory NLKEMIEAGVHFGHGKKNWPKMAPYISAKRKGTHITNLARTARFLSEACDLVFAASQ  
14pin DLEEMMEARVHLGHKTRKWNPKMAPYIYTERKDTIINLAKTARSLSEACDLFLIAGRG  
15por TLAELLEAGVHFGHQARRWNPMPFYIYTERNGIHIIDLVTQSLLTAECEVFKQASSDG  
16syn SLADLLESVGHFGHQTRWNPMDQYIYTERNGVHIIDLVTQALMEDAYEVRSSSTING

01can KRVLVGTKNQASKVVEREAKKCQ-----EFYINQRWPGGLTNWPTIKLSLKKLQLE  
02chl KTLFLVGTQQAAPLIAKTALACE-----SFYVNRWLGGMLTNWRTIQSLRKLQEYR  
03cya KKVLFIGTKRQAANI IAEQATRCG-----AFYINQRWGLLTNWTTIQSRVEYLKELE  
04eug KRALFVCTKRFFSILTQKIALKCN-----SFFVTKRWLGGILTNWITIKNCINKLKLS  
05gui KKFLLVGTQRQAASI IAEQAERCQ-----AFYINQRWGLLTNWTTIRTRVERLKDLE  
06mai KSFLIVGTKKRAADLVASAAIRSR-----CHYVNNKWFSGMLTNWSTIKTRLQSFDRDL  
07mar KQFLIVGTQYQAADLIESSALKAR-----CHYVNNKWLGGMLTNWSTIETRLQKFKDLE  
08mes KEFLVGTQKSQASDI IASEALRAE-----SHYVNRWLGGMLTNWSTIKTRIRKLKLE  
09nep GSFLVGTKNQAASI VRFAAERCQKNDRSQAHYVNRWLGGMLTNWPTMQLCIRLLELE  
10nic KQFLIVGTKNKAADSVEWAAIRAR-----CHYVNNKWLGGMLTNWSTIETRLHKFDRDL  
11odo QTFLFVGTQKQASAL IAEQAERCQ-----SYVNNRWLGGMLTNWVTLKSRIARLKTLE  
12oen GQFLIVGTQKQAAALVARAAIKAR-----GHYVNNKWLGGMLTNWSTIETRLHQFDRDL  
13ory KSFLIVGTKKRAADLVASAAIRAR-----CHYVNNKWFSGMLTNWSTIKTRLQSFDRDL  
14pin KQFLIVGTQYQATDLVASAATEAR-----CHYVNRWLGGMLTNWSTIETRLQKFKDLE  
15por RVLFLGTQKQAAAGI IAEQAERCQ-----SYVNNRWLGGMLTNWVTKSRVTRLRQLE  
16syn KKFLLVGTQRQAAGI ISQEAQRCG-----AHYVNRWLGGMLTNWETIRKRVDRLEKELE

01can EKEKLDHLGSTSKKEEASMKREIQKLNKYLNGIKGMKKPPDI IILVDPSEKTTISECQK  
02chl RAEERGDWNLKKQEVARKREKDRLEKYLSGVENMSRLPGVVILIGQTEEHAVKECRQ  
03cya MREESGALDLLPKKEAAILRRQLEKLRSKSGGIQTMRHIPDIVIIVDQKRENNAVQECK  
04eug KQKEK-HHNLLTKKERLVKKKKLKLKYYSGMRDMTERPEIIVIIIGQNKREINAVRECK  
05gui LKEETGYLDKLPKKEAAGLRELEKLKRNNGIKNMKRLPDLVVIIVDQKRETTAVQECRT  
06mai AEEKMGKFHHLPKRDAAILKRKLSTLQRYLGGIKYMTRLPDIIVILDQKQEIYALQECAI  
07mar NKKKTGTINRLPKKEAANLKRQLDHLQYLGKGYMTSLPDIIVIIDQKQEIYALQECIT  
08mes QKEKNGILDTLPKKEAANLKRQLYKRLKYLKYLQYLGKGYMTSLPDIIVIIDQKREINAVKECIT  
09nep SNHERGGFQFPKKEAAVARKQLARLEKHLGGVKNMTGLPTTVIIVGQKQEMNAVRECEK  
10nic MEQKTGRNLNRLPKRDAAMLKRQLSRLQYLGKGYMTGVPDIIVIIDQHEEYIALRECIT  
11odo QEEADQVFNLLPKKEASLRRKELEKLKRNNGIKDMEKLPDVAIIVDQKREMTAVTECRK  
12oen TEQKTGRNLNRLPKRDAAILKRQLSHLQYLGKGYMTGLPDIILIDQHEEYIALRECIT  
13ory AEEKMEKFHHLPKRDVAAILKRKLSTLQRYLGGIKYMTRLPDIIVILDQKQEIYALRECAI  
14pin KEQDTGRFNQLPKKEAAMLKRQLDQLQYLGKGYMTSLPDIIVIIDQKRETTAVQECRT  
15por EQDSSGLIDQLPKKEAANLRRLEKRLKHLNGIKNMTRLPDIIVVVDQKRETTAVQECLE  
16syn ALEASGGIDRRGKKEGSMRLRELKGLQYLGKGYMTSLPDIIVIIDQKRETTAVQECLE

01can LNIPTIGILDTNCPKSVNVP IPANDDSIKAIDLIVSKLSTSIKFAKLENNQFIENSTEN  
02chl LGIPTVTLTDSNCPNLADWFLPANDDSVSALRLILAWFQQAQIQTGLRLSLEEAACKQK  
03cya LNIPIIALLDTNSDPDVPDIIPGNDDAIRSIKLIIGKLADAILEGSHTKTSSAE-----  
04eug LGIASITILDTNCDPTLTYP IPSNDDSI LSVSLILSVLNSINRGVNNKVRQKFDKYK  
05gui LGIPIIISILDTNCPNLPDTIAIPGNDDAIRSIKLIITKLTDISEGSLGYDE-----NNS  
06mai LGIPTISLVDNCDPLANISIPANDDTMTSIRLILNKLVAISEGRSLYIRNR-----  
07mar LGIPTICLVDTDCTDPMTDIIPANDDARASIRVLNKLTLAICEGRYSIKN-----  
08mes LGIPTISLVDNCDPTLADLP IPANDDAIRSIINLVLKSLADAIEYEG--RVK-----  
09nep LGIPMITLDSNCPNLLTDIGIPSNDDSIASIELVMREITKGEIQQWIKQEGREATIE  
10nic LGIPTICLDTNCDPLADISIPANDDAISSIRLILNKLVAICEGRSSYIRNP-----  
11odo LGIPIIISILDTNCDPLVDIIPGNDDAVGSIKVLQSLTDSINTGKLTTK-----  
12oen LGIPTICLIDTDCDPLADLP IPANDDAMASIRLILNKLVAICEGRSSYIRNP-----  
13ory LGIPTISLADTNCNCPDLANISIPANDDTMTSIRLILNKLVAICEGRSLYIRNH-----  
14pin LGIPTICLVDTDCTDPLVDIIPANDDGIASIQILNRLTSAICEGRALRSL-----  
15por LGIPTICILDNCSPEIINIPANDDAIRSIKLIIGKIADAIYEGKYGMTEPTELISSA  
16syn LGIPIIAMLDNCPDVPDVP IPANDDAIRSIKLIIGKLADAIEYGRHGQPDVSDDYEEF

```

01can LNEHVK-----MGQKINPLSFRGLGINKLHSSWFARPQSY
02chl KKTGVKISGNRTSSITKKRNPASSKI---MGQKVHPIGFRLGITQKHSYWCCTTPQKS
03cya -----MGQKIHPIGFRGLGITQKHSWFANPQKY
04eug FKKLS-----MGQKVHPLGFRLGITKSHSSFWVVERRHY
05gui DHE-----MGQKVNPLGFRLRITSQHRSSWFATKESY
06mai -----MGQKINPLGFRLGTTQNHHSFWFAQPKNY
07mar -----MGQKINPLGFRLGITQNHRSWFAN-KKY
08mes -----MGQKIHPLGFRLGVTQEHRAHWFAKPSEY
09nep IEFVVEPTPVEKPLVIEPPPEKPVLSK-MGQKIHLPLGFRLGITQHRSTWFAPRKLY
10nic -----MGQKINPLGFRLGTTQGHSLWFSQPKNY
11odo -----VGQKTHPLGFRLGITQEHRSWAYANFNQY
12oen -----MAKINPLGFRLGTTQ-VIIRLVAQPKDY
13ory -----MGQKINPLGFRLGTTQNHHSFWFAQPKNY
14pin -----MAKINPLGFRLGVTQNHRSWFQAQQRNY
15por EDK-----MGQKIHPLGFRLGITQKHSWFASSKDY
16syn DEGLDGNLEVEAAEEVEAAEAEEVAATPEAMGQKIHVPFRLGITQKHSWYADPKRY

01can TAILQEDKKIRDYIFTNLLRA-----SISRQIN
02chl ALWIQDASFLRNFIKKYYIGA-----GITHIEIQ
03cya PTLQEDHLIRQYIEKNLSNA-----GIAQIYIQ
04eug ASFVKEDIVIRNFMNKELET-----LISLIKIE
05gui PQLLEQDFKIRSYINRELEAA-----GISKIEIS
06mai SEGLQEDKKIRNCKINYIQNRKKGSNRKMESDS-----SSEVITHIEIQ
07mar SKVFEEDKKIRDCIELYVQKH-----IKNSS-----NYGGIARVEIK
08mes KFLVEEDNYIRKYLNAKLANA-----GVARIDIQ
09nep VSWIHEERQLRDYLNKRLADA-----GVANVQLT
10nic SEGLQEDQKIRDCIKNYVQKN-----MRTSS-----GVEGIARIEIQ
11odo ANLLKEDNQIRTYLNKLAKTA-----SISNIQIN
12oen SDGLQEDQKIRDCIQNYIQKN-----MQISSDS-----GVEGIARIEIR
13ory SEGLQEDKKIRNCKINYIQNRKKGSNRKIEADSSFEVITHNKKMDSGSSSEVITHIEIQ
14pin SKDLREDQKIRTCIENYVRTH-----IKSSS-----NYGGIARVEIR
15por SVLLQEDHKIRSFHGLKLSNA-----SIKIEIN
16syn PELLQEDHKIRQYIEKTLNNA-----GISDIRIE

01can RQ-----FNQVELQLHTSRPGVVIIGRSG-TGIDSLKRNKVNLLSKQSQ-----
02chl RQDVDS-PLG---IQIYAARLRQIAGNDS-KGLERLQEELVKQLQIFYKRNLEVEPGTI
03cya RK-----ADRIEELRTARPGVVVGRGG-RGIEVLKRLKDLGEGQK-----
04eug RIYEFSE-QRNNITIVIHVARPERVIGRGG-QGLSRIRDILIDRMNYLLG-KT---PRII
05gui RN-----ANQLEVSVYTSRPGIIVGRSG-LGIEIKITDILRLKQDIS-----
06mai KE-----IDTIHVIIHIGFPNLLKKKG---AIEELEKDLQKEVNSVNO-----R--
07mar RK-----TDLIQVEIYTARPGLVVGRG-QGIEQLKLVNQNLSSDDR-----R--
08mes RK-----ADQVEIEIYTARPGLVVGRG-KGIEIIRDLQDKLRNKRK-----
09nep RQ-----SDRIEVEIHTACPGAIVGRG-QGLEILREDIQKRLPKVR-----
10nic KR-----IDLQVVIIFMGFPKLLIESRP-RGIEELQTTLQKEHFCVNR-----K--
11odo RN-----GLSDQIQLNIETGRPGVLVGENG-TGIKTLNLIKILPPNRQ-----
12oen KR-----IDLQVVIIFMGFPKLLIEDGT-RRIEELQRIVQKEINSVNR-----K--
13ory KE-----IDTIHVIIHIGFPNLLKKKG---AIEELEKDLQKEVNSVNO-----R--
14pin RK-----IDLKVKIYIGFPNLLIEGRGQIEKLNDVLNMLDSVDR-----K--
15por RK-----ADQVEVLIATARPGIVLKGSG-AGIESLRNSLTLILDPNKQ-----
16syn RK-----AEQIELGIHTARPGVVVGRGG-SGIEQLREGQLKLGSARQ-----

01can -LKINIIDVTNPDIADVLLACFISQQLESRTTFKRAVRQAIQRAQKS-DIPGIKIQVSGR
02chl HRLSVKPLRAPEAYAEVLAELVEELQKRPFRAMRQTQVQALRA-GVKGIVQVSGR
03cya -IRINIEVKQIDAEALIGEFITQQLERRVAFRIRVRAITRAQRR-GIEGKIQVSGR
04eug TCK--VVGVTSPNLDARLLADSVRRELEKRTPFIRAMKTVMQAMKA-GAEGIKVQVSGR
05gui -IRINIELTNPADANLIGEFIAQQLKRVAFRATRAQAIQKAQRA-NVQGIKQVSGR
06mai -LNIATIEKVKEPYRQPNILAEYIAFQLKNRVSFRKAMKKAIELTKKA-DIKGKIQIAGR
07mar -LRMTLIEIAKPYGEPKILAKKIALKLESRVAFRATMKAIELAKKG-NIKGKIQIAGR
08mes -FRITITTLQEPDLESTLIGEFIAQRLQERKPFRRATIRQAVQRAKRA-GVQGIKIQVAGR
09nep -VIVHVVEIANPDAQAVLIGGIIAKQLEERIPFRAVRQAVQRAMRAPGVEGKIEVSGR
10nic -LNIATVRIAKPYGNPNILAEYIAQQLKNRVSFRKAMKKAIELTEQA-DTGKIQIQIAGR
11odo -LTINIIEVEKVNLNASLIGDLVVEQLEDRVAFRAIRKAMQSALDE-NVNGIKIQVSGR
12oen -INITITKITKPYGDPNILAEYIAQQLKNRVSFRKAIQKAIELAEQA-DTGKIRVQIAGR
13ory -LNIATIEKVKEPYRQPNILAEYIAFQLKNRVSFRKAMKKAIELTKKT-DIKGVKVIAGR
14pin -LHIAIEKVAKPYRQPNILAEYIALQLEKRVFRTMKAIELAERE-EVEGIQIQIAGR
15por -IRNVVVEISDPDEATLVAEFITQQLKRVAFRAVRQAVQRAQRA-NTQGVKIQVSGR
16syn -IRNVVIEVPNADADAALMAEYIGQQLERRVSFRVVRQALQRAERA-EVKGKIQVSGR

01can LGAETIARTEWIREGRVPLQTLKADLDYATSSAYTSFGVVGVKWIIFKG-----
02chl LAGADIARSEDVREGVPLQTLRADIDYSSKPAKTFGLLGKIKIWFVRERLTRISNVR-
03cya LGAETIARSEWSREGVPLQTLRAEIDYSSKPAKTIYGVLGKVIWIFRGEVPGNPTEIS
04eug LAGIETIARTEWIREGRVPLHTLRADIDYFNDIAHTIYGVLGKIVWVYK-----
05gui LGAETIARSEWIREGRVPLQTLRANIDYATKEAHTTYGILGKIVWFNGEQTP-TYAVI-
06mai LAGKEIARAECIKKGRPLQTLIRAKIDYCCYPIRTIYGVLGKVIWIFVEE-----
07mar LGAETIARVEWAREGRVPLQTLIRANINCYAAQTIYGVLGKIVWIFQDEE-----
08mes LGAETIARSEWIREGRVPLQTLRADIDYSHCAKTIYGIIGKIVWIFRGEKISIN----
09nep LGAETIARSEWIREGRVPLHTLRADVDYCDCAQTIYGVLGKIVWIFRGEIRPVQVLAP
10nic LDGKEIARVEWIREGRVPLQTLIRAKIDYCSYTVRTIYGVLGKIVWIFLDEE-----
11odo LGAETIARSEWIREGRVPLQTLRADIDYATKEAHTIYGVLGKIVWIFKSEILAK-----
12oen INGGEIARVLWMKEGRVPLQTLIRAKIDYCYVTVRTIHGLLGKIVWIFRDE-----
13ory LAGKEIARAECIKKGRPLQTLIRAKIDYCCYPIRTIYGVLGKVIWIFVDEE-----
14pin LDGKEIARVEWDGRGRVPLQTLIRAKIDYCYPVKTIYGVLGKIVWIFRDE-----
15por LGAETIARSEWIREGRVPLQTLRADIDYCHRAHTTYGVLGKIVWIFRGEKELL-DSKVVE
16syn LGAETIARTEWIREGRVPLHTLRADIDYATRTALTYGILGKIVWIFRGEVPGQEAATV

```



```

01can -----MSRYTGPKIRIIRRLGELPALTTKKVK---NNYPPGR
02chl -----MARYRGPKLRIVRRLGELPGLTQKNCT---RDFPPGQ
03cya E-----MSRYKGPSLRIIRRLGELPGLTRKVVVK---RKYPGQ
04eug -----MSRYRGPLRLIRVRIGKLPSTLNKTSK---KRKSPGQ
05gui -----MSRYRGAVIKIIRRLGELPGLTRKTTT---RTSRPGQ
06mai -----MSRYRGPLKKIRRLGALPGLTRKTPKS---GSNQKKK
07mar -----MSRYRGPRVKIIRRLGALPGLTNKTLK---KSGYINQ
08mes -----MSRYRGPLKIVRKLGLPGLTSKINKNLQLAEQNKKG
09nep GT-----MSRYRGPLKIVRRLGELPGLTRKMAK---RKSPPGQ
10nic -----MSRYRGPRFKIRRLGALPGLTNKKPRN---GSDLRNQ
11odo -----MSRYRGPKLRIIRRLGALPGLTQKQSK---KKGRPGQ
12oen -----MSRYRGPRFKIRRLGALPGLTSKRPA---GSDPSRQ
13ory -----MSRYRGPRFKIRRLGALPGLTRKTPKS---GSNLKKK
14pin -----MSRYRGPLKIVRRLKTLPLGLTSKRPNK---RKDSMNR
15por LAPSQDQINPDVS-----MSRYRGPRVIRRLGDLPLSRKATK---RPYPGGE
16syn APPSQPRKRRQFDDRSQDMSRYRGPLRLIRVRRLGELPGLSRKSPR---RAYPPGQ

01can EWSTN-----EELSEYAIRLQEKQKIRFNNGINEQLRRYVKKAKSRGSTGSYLLNL
02chl HGPKKKGQGNQKTKESQYAVRLKEKQKLRFNNGISERQLMSYVREARKRGSTGEILLQI
03cya HGQKSR-----KRSEYAIRLEEKQKLRFNNGITERQLVQCVRAAKQMGSTGQILLQL
04eug PATSFK----RKKKISKYNIRLEEKQKLRFNNGITERQLLVYVKKSRKKGSSGRFLTLF
05gui HGTQAR-----KPSEYAIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGEALLQL
06mai FHSKKK-----EQYRIRLEEKQKLRFNNGITERQLLVYVTHAGAKRSTGQVLLQL
07mar STSNKKV-----SQYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL
08mes KSTKTKL-----SQYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL
09nep HGAASK-----KPSYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL
10nic SRSGKK-----SQYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL
11odo HGKSNE---ADNSKKTTEYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL
12oen GLSGNK-----SQYRIRLEEKQKLRFNNGITERQLLVYVTHAGAKRSTGQVLLQL
13ory FHSKKK-----EQYRIRLEEKQKLRFNNGITERQLLVYVTHAGAKRSTGQVLLQL
14pin S-SSRKI-----SQYRIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSRVRS-IAI
15por HGQKPR-----KPSEYAIRLEEKQKLRFNNGITERQLLVYVTHAGAKRKGSTGQVLLQL
16syn HGQARR-----KRSEYAIRLEEKQKLRFNNGITERQLLVYVTHAKRKGSTGQVLLQL

01can LEMRLDNIVLRAGLAPTIAASRQLVSHKHIEVNKNIVNPSFQCSIGDTIHVKNKNSRQ
02chl LEMRLDTIIFRLGFAPTIAAARQLINHGHIIVNGRRVDIPSSLCVNDNISVALN--SQN
03cya LEMRLDNIVFRLGMAPTIPASRQLVNHGHIIVNKNVVSIPSYQCPGDIITVKERNASKK
04eug LEMRLDNIVHRIGFAPTIAAKQLINHGHIIVNKNVVSIPSYQCPGDIITVKERNASKK
05gui LEMRLDNIVFRLGMAPTIPAAARQLVNHGHIIVNKNVVSIPSYQCPGDIITVKERNASKK
06mai LEMRLDNILFRLGMASTIPGARQLVNHHRILVNGRIVDIPSFRCPRDIITTKDNQSRK
07mar LEMRLDNIVFRLGMAPTIPGARQLVNHHRILVNGRIVDIPSFRCPRDIITTKDNQSRK
08mes LEMRLDNIVFRLGMAPTIPGARQLVNHHRILVNGRIVDIPSFRCPRDIITTKDNQSRK
09nep LEMRLDNIVFRLGMAPTIPGARQLVNHHRILVNGRIVDIPSFRCPRDIITTKDNQSRK
10nic LEMRLDNILFRLGMASTIPAAARQLVNHHRILVNGRIVDIPSYRCPKPRDIITAKDEQKSKA
11odo LEMRLDTICFTLGFAPTIAAARQLVNHGHIIVNKNVVSIPSYQCPGDIITVKERNASKK
12oen LEMRLDNILFRLGMASTIPGARQLVNHHRILVNGRIVDIPSYRCPKPRDIITAKDEQKSKA
13ory LEMRLDNILFRLGMASTIPGARQLVNHHRILVNGRIVDIPSFRCPRDIITTKDNQSRK
14pin TEMRLDKS-FSICYGSTS-GARQLVNHHRILVNDHMDVDIPSYCPKPDVITIRDQRLRA
15por LEMRLDNIVFRLGMAPTIPAAARQLVNHGHIIVNKNVVSIPSYQCPGDIITVKERNASKK
16syn LEMRLDNIVFRLGMAPTIPGARQLVCHGHITVNGQVVDIPSYQCPGDIIVSVDNRDRSRK

01can LIDLNSRRDITKFFPKYLEVNKDNMGARVVKTMKQDVNLITINELLVVEFYSRKG--MSQ
02chl FVKNLLQSFTKTLDAPLYEVNQEKL SAVVRDNP REAVSLQINELLVVEFYSRKV--MSR
03cya IVETNLLFPGLANLPSHLEFDKSKLQKQVNDIIQREWVALEVNELLIVEFYSR--MSR
04eug VIQKNIESKELLLIPPHLSLKNKLEAKIIGLINRKAISLVNELLVIEFYSRKV--MSR
05gui IVKNYLQFPGLANPNHLQIDKNDLTKINGI IERDQVVAL--NELLVIEFYSRKG--MSR
06mai LVQNYIASSDPGKLPKHLTVDTLQYKGLVKKILDRKWVGLKINELLVVEFYSRQT--MSR
07mar I IKNLNSFQKQIPNHLTFDLMQIKGLVNQI IREWIYLIKINELLVVEFYSRQV--MSR
08mes LVNLNLASPSVSKIPNHLTKKDTLTATVNGIVERKSIPLEINELLVVEFYSRQT--MSR
09nep LIEGYMSTMGSIPTPHLELKKEKLEGNQEIIDRQWVGLPINELLVVEFYSPKV--MSR
10nic LIQISLSDSSPHEELPNHLTHLPFPQYKGLVNQI IDSKWVGLKINELLVVEFYSRQT--MSR
11odo LVEGNLQTIKQVDLPTHLKFDKSKQEATVINYCDRNEILLNDELVLVIEYSSR--MSR
12oen MIQNSLESPPHEEVKHLTFHPFPQYKGLVNQI IDSEWVGLKINELLVVEFYSRQTKTMSR
13ory LVQNSIASSDPGKLPKHLTIDTLQYKGLVKKILDRKWVGLKINELLVVEFYSRQT--MSR
14pin I IKNIDLFQRDKLPNHLTFHSLQYKGLINQI IDSKWVGLKINELLVVEFYSRQA--MSR
15por LVESYLAFPGLANIPSHLELNKSNLQKNGVNDREWVALQINELLVVEFYSRK--MSR
16syn LVETNMEFPGLANVP SHLEFDKNTFTGKVNVSIDREWVALQINELLVIEFYSRKA--MSR

01can KDPYKTYRLVSDPFYESPLVTLIMHVLNKGKKSISQRIVYSAIENIAMVKEDPLEIIE
02chl RRTQKKRIVMPDPVYDSRLVELLVRQLMREGKKSLAYRICYESMNRVADATQDPLVIVE
03cya RSTAKKRILPDP IYNSRLVTLINHMLKDGKKSIA RSFIYEALKIIIEKKGSDPLEVLE
04eug RRRAKKR IISQDPIYNSTLASKVINIKLNGKKT LAQIF IYETMKNIQEIYKKDPLDILR
05gui RSTTKKKLALPDP IYNSRLVNMLTVRILKEGKKHLAQRIIYNAFDIIKQRTGEDAILVFE
06mai RGTAEKRTAKSDP IFRNRLVNMVNRIMKDGGKSLAYQILYRAVKKIQKKTETNPLVLRL
07mar KSTAEKQVAKPDP IYNSRLVNMLVNRILKNGKKS LAYQILYKAMKNIKQKTKNPLVLR
08mes RSTPKKRI IDSDPIYNSRLVMTLISHILKEGKSLAQKIFYQAMKNIEEKTEDPLKVLQ
09nep RNTAVKRSISSDPVYNSQLIHMISHILKEGKSLAYRLMYDAMKRIEKTQDQDILVVE
10nic RGTAEKRTAKSDP IYNSRLVNMLVNRILKNGKKS LAYQILYRAVKKIQKKTETNPLSVLR
11odo RNISKKRFP EADSTYNSYLVSLITRILKSGKKNLAQNI VNAAFEI IKVKTNE DPLVVE
12oen RGTAEKRTAKSDP IYNSRLVNMLINRILKNGKKS LAYQILYRAMKKIQKKTETNPLSVLR
13ory RGTAEKRTAKSDP IFRNRLVNMVNRIMKDGGKSLAYQILYRAVKKIQKKTETNPLVLRL
14pin RSTAEKRTAKSDP IYNSRLVNMLVNRILKNGKKS LAYQILYRAVKKIQKKTETNPLSVLR
15por RNTAKKRFA SPDP IYNSRLVMTLISHILKSGKKT LAQRIIYQALD IVKERTETDPLNVLE
16syn RGNVKKRPVPPDPVYNSLTL SMTIRRVMSRGGKSLASSIVYNALASVGKGTGEDPLEVFE

```

|       |                                                                |
|-------|----------------------------------------------------------------|
| 01can | RAIKNVIPAVEIRSRR-IGGSTYQVPTEVRVHRGISLSIRWVIFAKIRPGKSMCLKAN     |
| 02chl | QAIRNATPLVEVKARR-IGGSTYQVPLEVASERGTALAIRWLLSVCRKKTGRPMAAKLTA   |
| 03cya | QAVRNSTPLIEVKARR-IGGSTYQVPMVEVRDRTITLALRVVNSFSLQRLGKTIIVKLAN   |
| 04eug | KAIKNASPQMETRKR-IGGTIYQVPVEVKEDRGTSLALKFIEKARERKRGISTKLN       |
| 05gui | SAIKKVTPLEVKARR-IGGSTYQVPMVEVRAFRGTNLALRWITKYARERAGKSMCLKAN    |
| 06mai | QAIRRVTPNIGVKTTRNKKGSTRKVPMEIGSKQGRALAIRWLLASQKRPGRNMAFKLSS    |
| 07mar | QAVRKVTPNVTVKARR-IDGSTYQVPLEIKSTQKALAIRWLLGASRKRSQQNMFAKLSY    |
| 08mes | QAVLNATPLVEVKARR-LGGSTYQVPREVKAERGTAALRWLLSSARQRPGRNMFAKLTN    |
| 09nep | RAVRNATPTIEVKARR-MGGSTYQVPLEVKPERGTALALRWLLAARNRTGRDMVAKLSN    |
| 10nic | QAIRGVTPDITVKARR-VGGSTHQVPIEIGSTQKALAIRWLLAASRKRPRGRNMAFKLSS   |
| 11odo | RAIRNASPVVEVKARR-IGGSTYQVPEVSGFRATNLRLWIIQYSRQVRGRTMSIKLAN     |
| 12oen | QAIRRVTPNMTVKARR-VSGSTHPVPIEIGSTQGRALAIRWLLGASRKRPGRNMAFKLSS   |
| 13ory | QAIRRVTPNIGVKTTRNKKGSTRKVPMEIGSKQGRALAIRWLLASQKRPGRNMAFKLSS    |
| 14pin | QAIRRVTPNVTVKARR-VGGSTYRVPTIEIRSTQKVLAIRWLLGASRKRPGRNMFKLSH    |
| 15por | KAIRNITPLVEVKARR-VGGSTYQVPIEVRAVGTNLALRWITFRSERSGKSMCLKAN      |
| 16syn | KAIKNLTPLEVKARR-VGGATYQVPMVEVRARGTALALRWLVHFSRARGRTMESKLAN     |
|       |                                                                |
| 01can | ELLDASKNLGNSIRKKEDTHKMAEANRAFAHYRY---MVNDISDMLTRLNAICAQHKV     |
| 02chl | ELDDAAKNSGLAVRKRDEIHKMADANKAFAYRYF---MVNDTISDMLTRIRNANLAKKTS   |
| 03cya | ELIDAAANETGNTIKKREEMHRMAEANKAFVHYRY---MVNDTIADMLTRIRNANLAKHKV  |
| 04eug | EIIDASNNTGAEVKKKEEIHKTAEANKAFSNMKF---MTNIDVISDMLTRIRNLLIKARK   |
| 05gui | EIMDAANETGSSIRKKEETHRMAEANKAFAYHYR---VTNDTISDMLTRVRNANLAKHQV   |
| 06mai | ELVDAAGSGGAIKKKEATHRMAEANRALAHFR---MGKDTIADLLTSIRNADMNKKGT     |
| 07mar | ELIDAARDNGIAIRKKEETHKMAEANRAFAHFR---MGNDTIANMITSIRNANLKGKIKT   |
| 08mes | EIVDAANETGNAIRKKEETHRMAEANKAFVHYR---MVNDTIADMITRIRNANLITQKQ    |
| 09nep | ELMDASNRTIGNAVRKRDEMHRMAEANKAFAHIRV---MIQDTIADMLTRIRNANAMRIYT  |
| 10nic | ELVDAAGSGGDAIRKKEETHRMAEANRAFAHFR---MGRDTIAEIIITSIRNADMNRKRV   |
| 11odo | EIIDTANDIGNTIKKKEETHKMQMPKHLHIFVINY---MVTDTISDMLTRIRNANMIKHQI  |
| 12oen | ELVDATKGRGGAHTQKGRAHRMAEANRAFAHFR---MGRDTIADIITSIRNADMNRKGT    |
| 13ory | ELVDAAGGGGAIKKKEATHRMAEANRALAHFR---MGKDTIADLLTSIRNADMNKKGT     |
| 14pin | ELMDAARGNGNAIRKKEETHRMAEANRAFAHFR---MGNDTITNLITSIRNADMVEKGT    |
| 15por | EIMDAANETGNSIRKKEETHRMAEANKAFAYHYR---MVDNDTIVADMTRIRNANLARHQI  |
| 16syn | EIMDAANETGAAIKKREETHRMAEANKAFAYHYR---MASTDITISDMLTRIRNACAVRHST |
|       |                                                                |
| 01can | VQVPLTNINKNILKVLQKEGYIKNFE-----ILFERKSGYLLVSLKYNSLNQ-----      |
| 02chl | VSLPKTKVHEKMCQILEQEGFIKTF-----SFSETNTNELIVDLKYQDFASFGNYGV      |
| 03cya | ARVKATKITRCLANVLKEEGLIQNFE-----EIENNLQNELISLKYGKKR-----        |
| 04eug | VNVINTKLTVNIAEILKKEGFIDSFELADATCLTENGVIKKYITIFLKYKGPQK-----    |
| 05gui | VQVPATKMTKSIHVLLEEGFIESIE-----EVGLDINRQLLSLKYKGRER-----        |
| 06mai | VRVSTNITENIVKILLREGFIESVR-----KHQESNRYFLVSTLRHQRRKTRK---GI     |
| 07mar | VQVPATNITRNIKILFQEGFIDNFI-----DNKQNTKDILILNLKYQKKK---KS        |
| 08mes | VAVIASNTNKGIAQCILLKEGFIESIEY-----NTNSSNNPELILSLKYQKKR-----     |
| 09nep | VCMPMTSVAREIAVILETEGWID-----SWKEASVNSLILRLKYR-----G---AK       |
| 10nic | VRASTNITENIVQILLREGFIENVR-----KHREKNKYFLVTLRHRNRK---RP         |
| 11odo | VQIPATKMSKAITNILEEGFIEDYE-----IYMENSYQFLILSLKYKGSR-----        |
| 12oen | VRTESTNMPKEIVKILLREGFIENVR-----KHQENKKSFLVTLRHRNRKGP---SP      |
| 13ory | VRVSTNITENIVKILLREGFIESVR-----KHQESNRYFLVSTLRHQRRKTRK---GI     |
| 14pin | VRVTATNITKNIGRILLREGFIEDVR-----EHQEGQKYFLISTSKYRRRK---RT       |
| 15por | VQVPATKVTNRNALVLKEEGFVHNFE-----QMGEGETHLMISLKYNGKNR-----       |
| 16syn | TQVPTTKMTLSIAKVLKSEGFIEDYS-----ETGEGINKMLVLTLKYKGKTR-----      |
|       |                                                                |
| 01can | -DKPCLSVLKKISRPGLRMYVRTKKIPKVLGGTGIAIISTSKGVMTGSIAARNLGIGGEIL  |
| 02chl | G-KPCITNLKRISKPLRIYTNREIPKVLGGMGILILSTSKGLMTDRQARKLCLGGEIL     |
| 03cya | --QPIITALKRISKPLRLVYANHKEIPVLGGGLGIAIILSTSSGIMTDQTAHKGCGGEVL   |
| 04eug | --VSYITKIKRVSKPLRTYSSYKRLQSVAGGVGLTVLSTSKGLMTDLRLASNKIGGEIL    |
| 05gui | --EPVINALKRISRPGLRVYANKELPVLGGGLGIAIVISTSKGVLTDTKARTQGLGGEVL   |
| 06mai | YRTR---TFLKRISRPGLRIYANYQQIPKVLGGMGIAIILSTSRGIMTDREARLNRIGGEVL |
| 07mar | YIT---TLRRIKISRPGLRIYNSHKEIPKVLGGMGIVILSTSRGIMTDREARQKKIGGEIL  |
| 08mes | --KPYITALQRVSKGLRVYTSYKDIPKVLGGGIGIAIILSTSQGILTDQARMQKIGGEIL   |
| 09nep | Q-QPILTGLRRVSRSGCRVYSAKEVPKVLGGMGTAIISTSKGIMTDREARNHRLGGEVI    |
| 10nic | YRN---ILNLKRISRPGLRIYSNYQRIPIRLGGMGIVILSTSRGIMTDREARLEGIGGEIL  |
| 11odo | --EPVICKMVRVSKPLRVYSKSKLPKVLGNLIGIAIISTSKGVMTNLAKELGIGGEVL     |
| 12oen | SRTSLNLNLKRISRPGLRVYSNYQKIPRIPLGGMGIVILSTSRGIMTDREARLEGIGGEIL  |
| 13ory | YRTR---TFLKRISRPGLRIYANYQQIPKVLGGMGIAIILSTSRGIMTDREARLNRIGGEVL |
| 14pin | YMT---TSKRTSKPLRIYSNYREIPKVLGGMGIVILSTSQGILTDREARQKKIGGEIL     |
| 15por | --QPVITALKRISKPLRLVYANHKEIPVLGGGLGIAIISTSQGVMTDQARHDLGGEIL     |
| 16syn | --QPLINTLQRVSKPLRVYSPSKIPVLGGGIGIAIVSTSHGIMTDREARRQGIGGEIL     |
|       |                                                                |
| 01can | CYIW-----MKKIKRDKIVGAGKISNYLLLIIMLGGGISF                       |
| 02chl | CSVW-----MTAQELSIRYEVPGARRFGNYIWGSLMCLGGLGF                    |
| 03cya | CYIW-----MNLKNNTDQIKRDLITGSRRLSNYWAAITIGLSSGF                  |
| 04eug | FYIWMNLRDINMTLSKNENIKAKQKQINLPKILRQEIKNKIKIWFYINVMNLGGIGF      |
| 05gui | CYIW-----MNT-KIRTDLILGSKRFSNYAWCFILMTGGIGF                     |
| 06mai | CYIW-----MNRSEHIWIELLKGSRRKGNFFWACILFLGSLGF                    |
| 07mar | CYIW-----MNLQVDHIRVDFIIGSRRIISNFCWAFILLFGALGF                  |
| 08mes | CYIW-----MINFSTNSDLILRESVVGSRRIISNYWASVLLGASGF                 |
| 09nep | CLIW-----MDTTN--LVRRDIVIGSRRVSNYWASVLLGGSSF                    |
| 10nic | CYIW-----MTWRSEHIWIELITGSRKISNFCWAFILFLGSLGF                   |
| 11odo | CYIW-----MQKEIRRDDIIGSRRFSNYFWAVFLCSGGISF                      |
| 12oen | CYIW-----MTWRSEHIWIELIGSRKISNFCWAFILLFLGSLGF                   |
| 13ory | CYIW-----MNRSEHIWIELLKGSRRKGNFFWACILFLGSLGF                    |
| 14pin | CYIW-----MNRSKWLWIEPTIGSRKRSNFFWACILFLGSLGF                    |
| 15por | CYIW-----MKQVLPiHKVRKDVILGSRRFSNYWASTIFIGALGF                  |
| 16syn | CYIW-----MGGQTLAESSQVLQEVGLGARRFSNFFWAGISTIGGVGF               |

```

01can FIVGFLSYKFAFEHIALFSRFVNSDIRFIPQGITMIFYGTMAICLSIYIYFSIYYDIGAG
02chl LTIIGISSYLD----FPILSLVSSNIQFFPQGLVMCFYGLGFLLVGYIWLILWNLGEG
03cya LLAGISSYTK----INLLPFTDTTQFLFIPQGITMILLYGTIGFLLDIYLLWNLVNWGAG
04eug LIVGISSYIG----NMLIYFLDASEIIFFPQGITMCFYGTGCGILFSINQISILNGVGE
05gui CLTGVGSYFN----LHTILFVKFSDINFIPQGIWMMFYGTIAILFSLFLMYSIFTDVGGG
06mai LAVGASSYLK----KNMISVLPSSQILFFPQGVVMSFYGIAGLFISSYLWCTILWNVGSG
07mar FFVGFSSYLQ----KDLIPFLSAEQILFIPQGIWMMFYGIAGLFISFYLWCTICWNVGSG
08mes LIVGISSYLQ----YDLVPLSAKNIVFVPQGLVMCFYGSAGILLSIYWLTFWNVGEG
09nep LVVGLSSRLG----FDLVPLPAGDIIFIPQGLVMCFYGLVGLVSVSTYLVWTLWLSVGGG
10nic LLVGTSSYLG----RNLSIFFPQQIIFFPQGLVMSFYGIAGLFISSYLWCTISWNVGSG
11odo LLAGISSYFK----INFLPFANPKELAFIPQGLVMSFYGTLSIALAIYILGTLFWDIGSG
12oen LVVGTSSYLG----KNLISLVSSQIHFFPQGLVMSFYGIAGLFISSYMWCTISWNVGSG
13ory LAVGASSYLK----KNISVLPSSQILFFPQGVVMSFYGIAGLFISAYLWCTILWNVGSG
14pin FLVGISSYFG----ENLIPLLSSQILFVPQGIWMMFYGIAGLFISSYLWCTILFNVGSG
15por LLAGLSSYFQ----TDLLPFANSTELVFIPQGIWMTFYGSVGIFLSMFLWLTIIWNIGAG
16syn LLAGLSSYFQ----KNLLIVSDTTGLQFIPQGVALLFYGVAGSTVAGYLVWLTMLNVGSG

01can YNEFNLSKKVVVFRKGFPKGNRLIKFVLP IPHLKS VKVMARGG--INPKYEVFLYTKNLS
02chl YNEFNLETGYVRIFRWGFPKGNRRIDLQYPIQEIQSIRVEIQEG--INPKRIIYKLGRNR
03cya YNEYNKKKGTVSIFRWGFPKGNRRIEVIYPIEQIAIKLEIKQG--LNPRHSISLKIQEK
04eug YNEFNKELNMTIYRKKGKNSDINITYSKLDIEGIRIEIKNEYFNKQNVFLRIKDKN
05gui YNKYDKKEKEIEIFRLGYNKKNQMLKYNFDRDKSIKIELKDD--INPKREIYLVTKNKN
06mai YDRFDRKEGIVCIFRWGFPKGNRRIFLQFLVRDIQSIRIQVKEG--LYPRRIYMEIRGQG
07mar YNKFDKKGIFSIFRWGFPKGNRRIFLQFLVRDIQSIRMEVQEG--LPSRRVLYIKIKGQ
08mes YNEFDKVNGLVRIFRWGFPKGNRRINIVYDIKDIAIRVEIQEG--INPRRIYVLRKIGTR
09nep YNEFNKQEGVMRIFRWGFPKGNRRIFLQFLVRDIQSIRMEVQEG--MNPRRIYVLRKIGTR
10nic YDRFDRKEGIVCIFRWGFPKGNRRIFLQFLVRDIQSIRVEIQEG--ISARRVLYMDIRGQG
11odo YNEYNKVENLVKIVRKGFPGKNREILLTYPLTNIRAIGIKISEG--LNPKRSIYLCLEDR
12oen YDRFDRKEGIVCIFRWGFPKGNRRIFLQFLVRDIQSIRVEIQEG--IYARRVLYMEIKGQG
13ory YDRFDRKEGIVCIFRWGFPKGNRRIFLQFLVRDIQSIRIQVKEG--LPSRRVLYMEIRGQG
14pin YNKFDKKGIVCLFRWGFPKGNRRIFLQFLVRDIQSIRMEVQEG--ISARRVLYMEIKGRQ
15por YNEFNKNEGIVKIFRLGFPKGNRRIFLQFLVRDIQSIRMEVQEG--LNPRRIYVLRKIGTR
16syn YNEFNKKSQGVTIFRWGFPKGNRRIFLQFLVRDIQSIRMEVQEG--VNPKRSLYLVKQKRR

01can RPIQGVFKL----SKLEYQASEIATFLGLA-FEYNL-----
02chl EIPLTRAGQP--LSIQIETQAAELAKFLQVS-LEGI-----
03cya EIVITPIGYL--LPISVVEEQAAELASFLNIP-LDSNQ-----
04eug DLPIIQLSNP--TKISDLEKQASEIASFLNIP-IGY-----
05gui QIPLTRIGEP--LLSDVENQAEIELANFLNIP-IEGI-----
06mai VIPLTRTDEKFTPREIEQKAAELAYFLRVP-IEVF-----
07mar DIPLSRIEEY--FTLREMEDKAAELARFLKVS-IEGI-----
08mes DMPLTRIGQP--LTLAEIEEQAAELARFLQVN-IEGI-----
09nep EVPLTRIGQP--LTLAEIEQKAAELAGFLQVS-LEGF-----
10nic SIPLTRTDEN--LTLREIEQKAAELAYFLRVP-IEVF-----
11odo QIPLTPVQQP--NSISNLEEEAELAKFLDLK-LENL-----
12oen AIPLTRSDDN--LTLREIEQKAAELAYFLRVP-IEVF-----
13ory AIPLTRTDEKFTPREIEQKAAELAYFLRIP-MEVF-----
14pin DIPLTRTGDN--VNLREIEQKAAELARFLRVS-IEGF-----
15por NIPLTRVQQP--LLSEVEEQAAELARFLDVV-LEGA-----
16syn DIPLTRAGQP--ISLSQLENQAAELARFLGVP-LEGL--MCFPQGTYGSLLRERFNDNIC

01can -----MSILLQFFIISIIFFSLLILVILVPSQLSLQSGWQVSK
02chl -----MLLIQQLALFAIVVSFLLVVGVPVVLATPEGWAENK
03cya -----MLIAFQGAFAVFLVLSFVLIVAVPALASPGWERSQ
04eug -----MLLFTFTFQALVLALIFSFILVLTLPVIFASPGWENNK
05gui -----MTIILQLLVFALISFLLVIGVPLVSPDGWESSK
06mai -----MNIQFQALVFLVATSSVLVIRGHLVSPDGWSSNK
07mar -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
08mes -----MLAQLTFLVATSSVLVIRGHLVSPDGWSSNK
09nep -----MTIFQLALFALVATSSVLVIRGHLVSPDGWSSNK
10nic -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
11odo -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
12oen -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
13ory -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
14pin -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
15por -----MTIAFQLAVFALVATSSVLVIRGHLVSPDGWSSNK
16syn PYQLSRISAGDSLAQFIRRNKSMISVFGIALAALVLSFVVMVGPVATASPDGWDNRSK

01can SRFIALFTIWASMLISGFISVFV---
02chl STVFSGIGIWFLVFLVILNLSFV---
03cya RLIIYAGAAWLSLIIIVIGVLDVSVANQA
04eug SRIWLACRFWFLVFLVILNLSFV---
05gui NLVYASAGLWFLVIVTAAFNSFVI---
06mai NVVFSGTSLWIGLVLVAILNLSIS---
07mar NVVFSGASLWIGLVLVAILNLSIS---
08mes NVVFSGALLWISLVFAVGILNLSFV---
09nep GYVFGQVSAWFLVFTVGVNLSLVA---
10nic NVVFSGTSLWIGLVLVAILNLSIS---
11odo STFTRAFQAWVGLVIVIAAADGISSAI-
12oen NVVFSGTSLWIGLVLVAILNLSIS---
13ory NVVFSGTSLWIGLVLVAILNLSIS---
14pin NVVFSGVSGLWIGLVLVAILNLSIS---
15por GLIYTGAGLWGLVIVTSLVNLV---
16syn PLLYLGSGIWAIVVAILNLSLV---

```

# Literaturverzeichnis

- Adachi, J. and Hasegawa, M. (1996). *MOLPHY Version 2.3, Programs for Molecular Phylogenetics Based on Maximum Likelihood*. Insitute of Statistical Mathematics, Tokyo.
- Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Ponten, T., Alsmark, U. C., Podowski, R. M., Naslund, A. K., Eriksson, A. S., Winkler, H. H., and Kurland, C. G. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria . *Nature*, **396**, 133–140.
- Blattner, F. R., Plunkett, G., and III, Bloch, C. A. e. a. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1474.
- Bronstein, I. N. and Semendjajew, K. A. (1991). *Taschenbuch der Mathematik*. B. G. Teubner Verlagsgesellschaft, Stuttgart - Leipzig and Verlag Nauka, Moskau.
- Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. D., Geoghagen, N. S., Weidman, J. F., Fuhrmann, J. L., Nguyen, D. T., Utterback, T., Kelley, J. M., Peterson, J. D., Sadow, P. W., Hanna, M. C., Cotton, M. D., Hurst, M. A., Roberts, K. M., Kaine, B. B., Borodovsky, M., Klenk, H. P., Fraser, C. M., Smith, H. O., Woese, C. R., and Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* . *Science*, **273**, 1058–1073.
- Cavalier-Smith, T. (1998). A revised six-kingdom system of life. *Biol. Rev.*, **73**, 203–266.
- Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry III, C. E., Tekaiia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Sqaes, R., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence . *Nature*, **393**, 537–544.

- Corliss, J. O. (1994). An Intern Utilitarian (User friendly) Hierarchical Classification and Characterization of the Protists. *Acta Protozoologica*, **33**, 1–51.
- Darwin, C. (1859). *The Origin of Species by Means of Natural Selection*. Murray, London.
- Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., G. D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., H. R., Feldman, R. A., Short, J. M., Olson, G. J., and Swanson, R. V. (1998). The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature*, **392**, 353–358.
- Delwiche, C. (1999). Tracing the Thread of Plastid Diversity through the Tapestry of Life. *Am. Nat.*, **154**, 164–177.
- Dickerson, R. (1980). Evolution and gene transfer in purple photosynthetic bacteria. *Nature*, **283**, 210–212.
- Doolittle, W. F. (1999). Phylogenetic Classification and the Universal Tree. *Science*, **284**, 2124–2128.
- Douglas, S. and Penny, S. (1999). The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved syntenic groups confirm its common ancestry with red algae. *J. Mol. Evol.*, **48**, 236–244.
- Felsenstein, J. (1993). *PHYLIP (Phylogeny Interference Package) version 3.5c*. Department of Genetics, Washington State University Seattle.
- Figge, R. M. and Cerff, R. (2001). GAPDH gene diversity in spirochetes: A paradigm for genetic promiscuity. *Mol. Biol. Evol.*, **18**, 2240–2249.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G. G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, **269**, 496–512.
- Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G. G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J. L., Nguyen, D. T., Utterback, T., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J., Dougherty, B. A., Bott, K. F., Hu, P. C., Lucier, T. S., Peterson, S. N., Smith, H. O., and Venter, J. C. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science*, **270**, 397–403.

- Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R. A., L. R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., G. M., Dougherty, B., Tomb, J.-F., Fleischmann, R. D., R. D., Peterson, J., Kerlavage, A. R., Quackenbush, J., S. S., Hanson, M., van Vugt, R., Palmer, N., Adams, M. D., G. J. D., Weidman, J., Utterback, T., Wathley, L., McDonald, L., A. P., Bowman, C., Garland, S., Fujii, C., Cotton, M. D., Horst, K., R. K., Hatch, B., Smith, H. O., and Venter, J. C. (1997). Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. *Nature*, **390**, 580–586.
- Fraser, C. M., Norris, S. J., Weinstock, G. M., White, O., Sutton, G. G., Dodson, R., Gwinn, M., Hickey, E. K., Clayton, R., Ketchum, K. A., Sodergren, E., Hardham, J. M., McLeod, M. P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J. K., Chidambaram, M., Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M. D., Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Wathley, L., Weidman, J., Smith, H. O., and Venter, J. C. (1998). Complete Genome Sequence of *Treponema pallidum*, the Syphilis Spirochete. *Science*, **281**, 375–388.
- GCG (1994). *Program manual for Version 8*. Genetics Computer Group, 575 Sience Drive, Madison, Wisconsin, 53711, USA.
- Goksøyr, J. (1967). Evolution of eucaryotic cells. *Nature*, **214**, 1161.
- Gram, C. (1884). Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte der Medicin*, **2**, 185–189.
- Graur, D., W.-H. L. (1999). *Fundamentals of Molecular Evolution*. Sinauer Associates, Inc.
- Gupta, R. S. (1998). Protein Phylogenies and Signature Sequences: A Reappraisal of evolutionary Relationship among Archaeobacteria, Eubacteria and Eukaryotes. *Microbiol. Mol. Biol. Rev.*, **62**, 1435–1491.
- Haeckel, E. (1866). *Generelle Morphologie der Organismen. Zweiter Band: Allgemeine Entwicklungsgeschichte der Organismen*. Verlag Georg Reimer, Berlin.
- Hallick, R. B., L., H., Drager, R., M., F., and Monfort, A. Orsat, B. e. a. (1993). Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Res.*, **21**, 3537–3544.
- Hansmann, S. (1997). Einfluß der Variabilität von Aminosäurepositionen auf die Rekonstruktion phylogenetischer Bäume. Studienarbeit.
- Hansmann, S. (1998). Algorithmen zur Konstruktion phylogenetischer Bäume. Diplomarbeit.
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Umayam, L. A., Gill, S. R., Nelson, K. E., Read, T. D., Tettelin, H., Richardson, D., Ermolaeva, M. D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P.,



- McDonald, L., Utterback, T., Fleishmann, R. D., Nierman, W. C., White, O., Salzberg, S. L., Smith, H. O., Colwell, R. R., Mekalanos, J. J., Venter, J. C., and Fraser, C. M. (2000). DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477–483.
- Hendy, M. and Penny, D. (1993). Spectral Analysis of Phylogenetic Data. *J. Classification*, **10**, 5–24.
- Himmelreich, R., Hilbert, H., Plagens, H., Pirkel, E., Li, B. C., and Herrmann, R. (1996). Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.*, **24**, 4420–4449.
- Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., and Sakamoto, M. Mori, M. e. a. (1989). The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.*, **217**, 185–194.
- Hupfer, H., Swiatek, M., Hornung, S., Herrmann, R. G., Maier, R. M., Chiu, W. L., and Sears, B. (2000). Complete nucleotide sequence of the *Oenothera elata* plastid chromosome, representing plastome I of the five distinguishable *Euoenothera* plastomes. *Mol. Gen. Genet.*, **263**, 581–585.
- Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W., O. L., Grimwood, J., Davis, R., and Stephens, R. (1999). Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nat. Genet.*, **21**, 385–389.
- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpō, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, **3**, 109–136.
- Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Yoshizawa, T., Nakamura, Y., Robb, F. T., Horikoshi, K., Masuchi, Y., Shizuya, H., and Kikuchi, H. (1998). Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3 (supplement). *DNA Res.*, **5**, 147–155.
- Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., J.-n. K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Nakazawa, H., T.-M., Masuda, S., Funahashi, T., Tanaka, T., Kudoh, Y., Y.-J., Kushida, N., Oguchi, A., Aoki, K.,

- Kubota, K., N.-Y., Nomura, N., Sako, Y., and Kikuchi, H. (1999). Complete Genome Sequence of an Aerobic Hyper-thermophilic Crenarchaeon, *Aeropyrum pernix* K1 . *DNA Res.*, **6**, 83–101.
- Kishino, H. and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.*, **29**, 170–179.
- Klenk, H. P., Clayton, R. A., Tomb, J., White, O., Nelson, K. E., K. K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., R. D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., F. R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., K. E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B., P. S., Reich, C. I., McNeil, L. K., Badger, J. H., Glodek, A., Z. L., Overbeek, R., Gocayne, J. D., Weidman, J. F., McDonald, L., U. T., Cotton, M. D., Spriggs, T., Artiach, P., Kaine, B. P., S. S. M., Sadow, P. W., D'Andrea, K. P., Bowman, C., Fujii, C., G. S. A., Mason, T. M., Olsen, G. J., Fraser, C. M., Smith, H. O., W. C. R., and Venter, J. C. (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus* . *Nature*, **390**, 364–370.
- Kowallik, K., Stoebe, B., Schaffram, I., Kroth-Pancic, P., and Freier, U. (1995). The chloroplast genome of a chlorophyll a+c containing alga, *Odontella sinensis*. *Plant Mol. Biol. Rep.*, **13**, 336–342.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., A. V., Berto, M. G., Bessieres, P., Bolotin, A., Borchert, S., B. R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., B. S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., C. N. M., Choi, S. K., Codani, J. J., Connerton, I. F., and Danchin, A. et. a. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis* . *Nature*, **390**, 249–256.
- Lemieux, C., Otis, C., and Turmel, M. (2000). Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature*, **403**, 649–652.
- Lockhart, P. J., Steel, M. A., Hendy, M. D., and Penny, D. (1994). Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.*, **11**, 605–612.
- Lozan, J. and Kausch, H. (1998). *Angewandte Statistik für Naturwissenschaftler*. Parey Buchverlag Berlin.
- Lyons-Weiler, J. and Hoelzer, G. A. (1996). Relative Apparent Synapomorphy Analysis (RASA) I: The Statistical Measurement of Phylogenetic Signal. *Mol. Biol. Evol.*, **13**, 749–757.
- Lyons-Weiler, J. and Hoelzer, G. A. (1997). Escaping from the Felsenstein Zone by detecting Long Branches in Phylogenetic Data. *Mol. Phylog. Evol.*, **8**, 375–384.



- Lyons-Weiler, J. and Hoelzer, G. A. (1999). Null Model Selection, Compositional Bias, Character State Bias, and the Limits of Phylogenetic Information. *Mol. Biol. Evol.*, **16**, 1400–14005.
- Maier, R., Neckermann, K., G., I., and Kössel, H. (1995). Complete sequence of maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J. Mol. Biol.*, **251**, 614–628.
- Martin, W. and Hansmann, S. (2000). Phylogeny of 33 ribosomal and six other proteins encoded in an ancient, conserved gene cluster that is conserved across prokaryotic genomes: influence of excluding poorly alignable sites from analysis. *Int. J. Syst. Evol. Microbiol.*, **50**, 1655–1663.
- Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., and Kowallik, V. (1998). Gene Transfer to the nucleus and the evolution of chloroplasts. *Nature*, **393**, 162–165.
- Martin, W., Hoffmeister, M., Rotte, C., and Henze, K. (2001). An Overview of Endosymbiotic Models for the Origins of Eukaryotes, Their ATP-Producing Organelles (Mitochondria and Hydrogenomes), and Their Heterotrophic Lifestyle. *Biol. Chem.*, **382**, 1521–1539.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press.
- Nei, M. (1996). Phylogenetic analysis in molecular evolutionary genetics. *Ann. Rev. Genet.*, **30**, 371–403.
- Nei, M. and Saitu, N. (1987). The Neighbor-Joining-Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.*, **4**, 406–425.
- Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999). Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature*, **399**, 323–329.
- Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organisation deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature*, **322**, 572–574.
- Pace, N. (1997). A molecular view of microbial diversity and biosphere. *Science*, **276**, 734–730.
- Parkhill, J., Wren, B., Mungall, K., Ketley, J., Churcher, C., B. D., Chillingworth, T., Davies, R., Feltwell, T., Holroyd, S., J. K., Karleyshev, A., Moule, S., Pallen, M., Penn, C.W., Q. M., Rajandream, M., Rutherford, K., VanVliet,

- A., W. S., and Barrell, B. (2000). The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature*, **403**, 665–668.
- Penny, D. and Hendy, M. D. (2001). Phylogenetics and distance methods. In Blading, D.J., Bishop, M. and Cannings, C., editor, *Statistical Genetics*, pages 445–484. John Wiley and Sons.
- Reith, M. and Mumholland, J. (1995). Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol. Biol. Rep.*, **13**, 333–335.
- Sagan, L. (1967). On the origin of mitosing cells. *J. Theoret. Biol.*, **14**, 225–274.
- Schlegel, H. G. (1992). *Allgemeine Mikrobiologie*. Thieme Verlag.
- Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., and Matsumayashi, T. e. a. (1986). The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression. *EMBO J.*, **5**, 2043–2049.
- Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C. R., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorri, H., Facincani, A. P., Ferreira, A. J. S., Ferreira, V. C. A., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H. S., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C. C., Lemos, E. G. M., Lemos, M. V. F., Lopes, S. A. Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A. M. B. N., Madeira, H. M. F., Marino, C. L., Marques, M. V., Martins, E. A. L., Martins, E. M. F., Matsukuma, A. Y., Menck, C. F. M., Miracca, E. C., Miyaki, C. Y., Monteiro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A. L. T. O., Netto, L. E. S., Jr., N., A., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A. G., Jr., P., A., H., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., de M. Rosa, A. J., de Rosa Jr., E., V., de Sa, R. G., Santelli, R. V., S. H. E., da Silva, A. C. R., da Silva, F. R., da Silva, A. M., Jr., S., A., W., da Silveira, J. F., Silvestri, M. L. Z., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., and Setubal, J. C. (2000). The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature*, **406**, 151–157.
- Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H.-M., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison,

- D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicare, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Safer, H., Patwell, D., Prabhakar, S., McDougall, S., Shimer, G., Goyal, A., Pietrovski, S., Church, G. M., Daniels, C. J., Mao, J.-I., Rice, P., Nolling, J., and Reeve, J. N. (1997). Complete genome sequence of *Methanobacterium thermoautotrophicum* delta H: functional analysis and comparative genomics. *J. Bacteriol.*, **179**, 7135–7155.
- Stainer, R. Y. and van Niel, B. C. (1941). The main outlines of bacterial classification. *Bacteriol.*, **42**, 437–466.
- Steel, M. A. (1994). Recovering a tree from the leaf colourations it generates under a Markov model. *Appl. Math.*, **7**, 19–24.
- Stephens, R. S., Kalman, S., Lammel, C. J., Fan, J., Marathe, R., A. L., Mitchell, W. P., Olinger, L., Tatusov, R. L., Zhao, Q., K. E. V., and Davis, R. W. (1998). Genome Sequence of an Obligate Intracellular Pathogen of Humans: *Chlamydia trachomatis*. *Science*, **282**, 754–759.
- Stirewalt, V., Michalowski, C., Löffelhardt, W., Bohnert, H., and Bryant, D. (1995). Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. *Plant Mol. Biol. Rep.*, **13**, 327–332.
- Stoebe, B., Hansmann, S., Goremykin, V., Kowalik, K., and Martin, W. (1999). Proteins encoded in sequenced chloroplast genomes : An overview of gene content, phylogenetic information, and endosymbiotic gene transfer to the nucleus. In Hollingsworth C., Batemann R. and Gornall M., editors, *Advances in Plant Molecular Systematics*, pages 327–352. Francis and Taylor, Andover.
- Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltry, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., and Paulsen, I. T. (2000). Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, **406**, 959–964.
- Tettelin, H., Saunders, N. J., Heidelberg, J., Jeffries, A. C., Nelson, K. E., Eisen, J. A., Ketchum, K. A., Hood, D. W., Peden, J. F., Dodson, R. J., Nelson, W. C., Gwinn, M. L., DeBoy, R., Peterson, J. D., Hickey, E. K., Haft, D. H., Salzberg, S. L., White, O., Fleischmann, R. D., Dougherty, B. A., Mason, T., Ciecko, A., Parksey, D. S., Blair, E., Cittone, H., Clark, E. B., Cotton, M. D., Utterback, T. R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Massignani, V., Pizza, M., Grandi, G., Sun, L., Smith, H. O., Fraser, C. M., Moxon, E. R., Rappuoli, R., and Venter, J. C. (2000). Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science*, **287**, 1809–1815.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through

- sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
- Tomb, J.-F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzgerald, L. M., Lee, N., Adams, M. D., Hickey, E. K., Berg, D. E., Gocayne, J. D., Utterback, T. R., Peterson, J. D., Kelley, J. M., Karp, P. D., Smith, H. O., Fraser, C. M., and Venter, J. C. (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature*, **388**, 539–547.
- Turnel, M., Otis, C., and Lemieux, C. (1999). The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: Insights into the architecture of ancestral chloroplast genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 10248–10253.
- van Niel, C. B. (1946). The classification and natural relationship of bacteria. *Quant. Biol.*, **11**, 285–301.
- Vogel, H., Fischer, S., and Valentin, K. (1996). A model for the evolution of the plastid sec apparatus inferred from secY gene phylogeny. *Plant Mol. Biol.*, **32**, 685–692.
- Wächtershäuser, G. (1998). Towards a reconstruction of ancestral genomes by gene cluster Alignment. *Syst. Appl. Microbiol.*, **21**, 473–477.
- Wakasugi, T., T., N., Kapoor, M., Sugita, M., Ito, M., and Ito, S. e. a. (1997). Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 5967–5972.
- Wakasugi, T., T. J., Ito, S., Nakashima, K., Tsudzuki, T., and M., S. (1994). Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 9794–9798.
- White, O., Eisen, J. A., Heidelberg, J. F., Hickey, E. K., Peterson, J. D., D. R. J., Haft, D. H., Gwinn, M. L., Nelson, W. C., Richardson, D. L., M. K. S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., V. J. J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., M. K. S., Aravind, L., Daly, M. J., Minton, K. W., F. R. D., Ketchum, K. A., Nelson, K. E., Salzberg, S., S. H. O., Venter, J. C., and Fraser, C. M. (1999). Genome Sequence of the Radioresistant Bacterium *Deinococcus radiodurans* R1. *Science*, **286**, 1571–1577.
- Woese, C. R. (1987). Bacterial Evolution. *Microbiol. Rev.*, **51**, 221–271.
- Woese, C. R. and Fox, G. E. (1977). Phylogenetic structure of the procaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA*, **95**, 5088–5090.

- Zuckerkandl, E. and Pauling, L. (1965). Molecules as documents of the evolutionary history. *J. Theor. Biol.*, **8**, 357–366.

## Lebenslauf

- **Name:** Sabine Cornelsen geb. Hansmann
- **Geburtsdatum:** 19.9.1972
- **Geburtsort:** Braunschweig
- **Familienstand:** verheiratet
- **Wohnort:** Braunschweig
- **Eltern:** Jutta und Dieter Hansmann
- **schulischer Bildungsweg (1979 - 1992):**
  - 1979-1983: Grundschule Braunschweig-Gliesmarode
  - 1983-1985: Orientierungsstufe Braunschweig-Querum
  - 1985-1992: Gymnasium Neue Oberschule Braunschweig
  - Schulabschluß: Abitur
- **Studium (1992 - 1998):**
  - Immatrikulation: 1.10.1992 im Studiengang Informatik
  - Vordiplom: 3.3.1995
  - Diplom: 30.9.1998

Meine Diplomprüfungen habe ich in den Fächern Compilerbau, Robotik, Theoretische Informatik und dem Anwendungsfach Genetik abgelegt. In meiner Studien- und Diplomarbeit behandelte ich Problemstellungen aus dem Bereich der Bioinformatik mit dem Themenschwerpunkt der Evaluierung und Untersuchung von Algorithmen zu Rekonstruktion phylogenetischer Abstammungsbäume.

- **Promotionszeit:**
  - seit 1.4.1999 Promotionsstipendiatin der Heinrich Böll Stiftung
  - Promotionsthema: Bioinformatische Untersuchungen zur Evolution der Prokaryoten und Plastiden
  - Betreuung: Prof. W.F.Martin, Institut für Botanik III, Heinrich Heine Universität Düsseldorf
  - November 2000 - Januar 2001 Forschungsaufenthalt am Institute of Molecular Bioscience, Massey University, Palmerston North, Neuseeland
- **sonstige berufliche Aktivitäten:**
  - seit Beginn meiner Studienarbeit 1996 zeitweilige Beschäftigungen als studentische bzw. wissenschaftliche Hilfskraft am Institut für Genetik